Cochrane Library

Cochrane Database of Systematic Reviews

Paying for performance to improve the delivery of health interventions in low- and middle-income countries (Review)

Witter S, Fretheim A, Kessy FL, Lindahl AK

Witter S, Fretheim A, Kessy FL, Lindahl AK.
Paying for performance to improve the delivery of health interventions in low- and middle-income countries.
Cochrane Database of Systematic Reviews 2012, Issue 2. Art. No.: CD007899.
DOI: 10.1002/14651858.CD007899.pub2.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 3
BACKGROUND 5
OBJECTIVES 7
METHODS 7
RESULTS 10
DISCUSSION 13
AUTHORS' CONCLUSIONS 15
ACKNOWLEDGEMENTS 16
REFERENCES 17
CHARACTERISTICS OF STUDIES 19
ADDITIONAL TABLES 30
APPENDICES 60
WHAT'S NEW 69
HISTORY 69
CONTRIBUTIONS OF AUTHORS 70
DECLARATIONS OF INTEREST 70
SOURCES OF SUPPORT 70
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 70
INDEX TERMS 70

[Intervention Review]

Paying for performance to improve the delivery of health interventions in low- and middle-income countries

Sophie Witter[1], Atle Fretheim[2,3], Flora L Kessy[4], Anne Karin Lindahl[5]

[1]Immpact, University of Aberdeen, Aberdeen, UK. [2]Global Health Unit, Norwegian Knowledge Centre for the Health Services, Oslo, Norway. [3]Institute of Health and Society, University of Oslo, Oslo, Norway. [4]Ifakara Health Institute (IHI), Dar es Salaam, Tanzania. [5]Norwegian Knowledge Centre for the Health Services, Oslo, Norway

Contact address: Sophie Witter, Immpact, University of Aberdeen, Health Sciences Building, Foresterhill, Aberdeen, AB25 2DZ, UK. sophiewitter@blueyonder.co.uk.

Editorial group: Cochrane Effective Practice and Organisation of Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 5, 2013.

Citation: Witter S, Fretheim A, Kessy FL, Lindahl AK. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database of Systematic Reviews* 2012, Issue 2. Art. No.: CD007899. DOI: 10.1002/14651858.CD007899.pub2.

ABSTRACT

Background

There is a growing interest in paying for performance as a means to align the incentives of health workers and health providers with public health goals. However, there is currently a lack of rigorous evidence on the effectiveness of these strategies in improving health care and health, particularly in low- and middle-income countries. Moreover, paying for performance is a complex intervention with uncertain benefits and potential harms. A review of evidence on effectiveness is therefore timely, especially as this is an area of growing interest for funders and governments.

Objectives

To assess the current evidence for the effects of paying for performance on the provision of health care and health outcomes in low- and middle-income countries.

Search methods

We searched more than 15 databases in 2009, including the Cochrane Effective Practice and Organisation of Care Group Specialised Register (searched 3 March 2009), CENTRAL (2009, Issue 1) (searched 3 March 2009), MEDLINE, Ovid (1948 to present) (searched 24 June 2011), EMBASE, Ovid (1980 to 2009 Week 09) (searched 2 March 2009), EconLit, Ovid (1969 to February 2009) (searched 5 March 2009), as well as the Social Sciences Citation Index, ISI Web of Science (1975 to present) (searched 8 September 2010). We also searched the websites and online resources of numerous international agencies, organisations and universities to find relevant grey literature and contacted experts in the field. We carried out an updated search on the Results-Based Financing website in April 2011, and re-ran the MEDLINE search in June 2011.

Selection criteria

Pay for performance refers to the transfer of money or material goods conditional on taking a measurable action or achieving a predetermined performance target. To be included, a study had to report at least one of the following outcomes: changes in targeted measures of provider performance, such as the delivery or utilisation of healthcare services, or patient outcomes, unintended effects and/or changes in resource use. Studies also needed to use one of the following study designs: randomised trial, non-randomised trial, controlled before-after study or interrupted time series study, and had to have been conducted in low- or middle-income countries (as defined by the World Bank).

Data collection and analysis

We aimed to present a meta-analysis of results. However, due to the limited number of studies in each category, the diversity of intervention designs and study methods, as well as important contextual differences, we present a narrative synthesis with separate results from each study.

Main results

Nine studies were included in the review: one randomised trial, six controlled before-after studies and two interrupted time series studies (or studies which could be re-analysed as such). The interventions were varied: one used target payments linked to quality of care (in the Philippines). Two used target payments linked to coverage indicators (in Tanzania and Zambia). Three used conditional cash transfers, modified by quality measurements (in Rwanda, Burundi and the Democratic Republic of Congo). Two used conditional cash transfers without quality measures (in Rwanda and Vietnam). One used a mix of conditional cash transfers and target payments (China). Targeted services also varied. Most of the interventions used a wide range of targets covering inpatient, outpatient and preventive care, including a strong emphasis on services for women and children. However, one focused specifically on tuberculosis (the main outcome measure was cases detected); one on hospital revenues; and one on improved treatment of common illnesses in under-sixes. Participants were in most cases in a mix of public and faith-based facilities (dispensaries, health posts, health centres and hospitals), though districts were also involved and in one case payments were made direct to individual private practitioners.

One study was considered to have low risk of bias and one a moderate risk of bias. The other seven studies had a high risk of bias. Only one study included any patient health indicators. Of the four outcome measures, two showed significant improvement for the intervention group (wasting and self reported health by parents of the under-fives), while two showed no significant difference (being C-reactive protein (CRP)-negative and not anaemic). The two more robust studies both found mixed results – gains for some indicators but no improvement for others. Almost all dimensions of potential impact remain under-studied, including intended and unintended impact on health outcomes, equity, organisational change, user payments and satisfaction, resource use and staff satisfaction.

Authors' conclusions

The current evidence base is too weak to draw general conclusions; more robust and also comprehensive studies are needed. Performance-based funding is not a uniform intervention, but rather a range of approaches. Its effects depend on the interaction of several variables, including the design of the intervention (e.g. who receives payments, the magnitude of the incentives, the targets and how they are measured), the amount of additional funding, other ancillary components such as technical support, and contextual factors, including the organisational context in which it is implemented.

PLAIN LANGUAGE SUMMARY

Paying for performance to improve the delivery of health interventions in low- and middle-income countries

Researchers in the Cochrane Collaboration conducted a review of the effect of paying for performance to improve the delivery of health care in low- and middle-income countries. After searching for all relevant studies, they found nine studies that met their requirements. The findings of this review are summarised below.

What is pay for performance?

In a 'pay for performance' approach, people are given money or other rewards if they carry out a particular task or hit a particular target. This approach can be directed at patients, health workers or healthcare organisations. Patients are sometimes rewarded if they use particular healthcare services. Health workers and healthcare organisations may be rewarded if they offer particular services, if they deliver care that is of a certain quality, or if their patients make use of particular services. This review focused on pay for performance approaches that target the behaviour of health workers and healthcare organisations.

What happens when health workers and healthcare organisations are paid for performance?

The quality of the evidence was generally very low. The pay for performance approaches used in each study varied a great deal and the studies were carried out in a wide range of settings. It is therefore not possible to draw general conclusions. There is a need for more and better research in this area.

SUMMARY OF FINDINGS

Summary of findings for the main comparison.

Patients or population: Providers of healthcare services in low- and middle-income countries

Settings: Vietnam, China, Uganda, Rwanda, Tanzania, Democratic Republic of Congo, Burundi, Philippines

Intervention: Performance-based financing (PBF)

Comparison: No performance-based financing

Outcomes	Impacts	Number of studies	Quality of the evidence (GRADE)
Provider performance (quality of care)	The impact of performance-based financing on service delivery is highly uncertain. Four studies measured coverage of tetanus vaccinations among pregnant women, with mixed findings. Results from one study showed a small or no impact on tuberculosis case detection.	5	⊕⊖⊖⊖ **Very low[1]**
Utilisation of services: antenatal care	The impact of performance-based financing on attendance rates for antenatal care is highly uncertain. The study results point in both negative and positive directions.	2	⊕⊖⊖⊖ **Very low[1]**
Utilisation of services: institutional deliveries	Whether performance-based financing leads to an increase in institutional deliveries is unclear. A wide range of effect estimates are reported in the studies, including substantially larger increases in areas *without* PBF, to almost a 2-fold increase in areas *with* PBF.	4	⊕⊖⊖⊖ **Very low[1]**
Utilisation of services: preventive care for children, including vaccination	Performance-based financing may or may not lead to increased utilisation of preventive care services for children. One study found that attendance rates for children's preventive services doubled, but the impact on immunisation rates ranged from negative to positive across the 4 studies.	4	⊕⊖⊖⊖ **Very low[1]**
Utilisation of services: number of outpatients	Utilisation of services may increase as a consequence of PBF, but this has not been rigorously evaluated and the studies where this has been assessed have not yielded consistent results.	4	⊕⊖⊖⊖ **Very low[1]**
Patient outcomes	The impact of performance-based financing on patient outcomes was evaluated in only 1 study. The results were inconsistent across the 4 measures that were used in the study: performance-based financing seemed to have an impact on rates of wasting and General Self Reported Health in this study, but not on CRP levels or on anaemia rates.	1	⊕⊕⊖⊖ **Low[2]**
Unintended effects	Only 2 studies reported on unintended effects – in both studies the authors voiced concerns about the curative nature of the coverage targets and whether this may squeeze out preventive care. However, no conclusive evidence was found to support or refute this.	2	⊕⊖⊖⊖ **Very low[3]**

Resource use	PBF payments tend to increase facility revenues and to increase staff pay, but their impact on wider resource use indicators, such as other funding sources, patient payments and efficiency of service provision are not yet established	8	⊕⊖⊖⊖ **Very low**[4]

*GRADE Working Group grades of evidence

⊕⊕⊕⊕ **High:** We are confident that the true effect lies close to what was found in the research.

⊕⊕⊕⊖ **Moderate:** The true effect is likely to be close to what was found, but there is a possibility that it is substantially different.

⊕⊕⊖⊖ **Low:** The true effect may be substantially different from what was found.

⊕⊖⊖⊖ **Very low:** We are very uncertain about the effect.

1.The evidence was by default graded as low as all studies were classified as non-randomised, and further downgraded to very low due to the high risk of bias and inconsistency across findings.
2.The evidence was downgraded to low because it was based on a singly study only, and because we considered several of the outcomes to be indirect (surrogate) measures of clinically important health outcomes.
3.The evidence was graded as very low as both studies were non-randomised with a high risk of bias, and because the outcome (unintended effects) was not thoroughly addressed in the studies.
4.The evidence was graded as very low as due to risk of bias in many of the studies, the lack of consistency in results, and the lack of reporting of impact on resource use indicators, independently of the intervention itself.

BACKGROUND

Description of the condition

Improving the performance of healthcare delivery systems is an important objective, both in high-income settings and, even more critically, in low- and middle-income country (LMIC) settings, where resources for health are much more constrained.

Performance-based financing (PBF) is currently receiving increased attention as a strategy for improving the performance of healthcare providers, organisations and governments. It is also promoted as an important tool for achieving the health Millennium Development Goals, improving the effectiveness of development aid, and motivating patients to improve their attendance at health facilities and compliance with recommended health interventions. However, there is currently a lack of rigorous evidence on the effectiveness of these strategies in improving health care and health, particularly in lower-income countries (Eldridge 2009; Oxman 2008).

Description of the intervention

Pay for performance refers to the transfer of money or material goods conditional on taking a measurable action or achieving a predetermined performance target (Eichler 2006). Paying for performance is also sometimes referred to as results-based financing, performance-based funding and output-based aid. While paying for performance is a relatively simple concept, it includes a wide range of interventions that vary with respect to the level at which the incentives are targeted (recipients of health care, individual providers of health care, healthcare facilities, private sector organisations, public sector organisations and national or sub-national levels). Paying for performance interventions can also reward a wide range of measurable actions, including health outcomes, delivery of effective interventions (for instance immunisation), utilisation of services (such as prenatal visits or births at an accredited facility) and quality of care. Paying for performance interventions typically also include ancillary components such as increasing the availability of resources to health care, education, supplies, technical support or training, monitoring and feedback, increasing salaries, construction of new facilities, improvements in planning and management or information systems etc (Oxman 2008).

While it is conceivable that pay increases designed to increase motivation and retention of staff might fall within this definition, in this review we focus on reforms which are explicitly linked to changing patterns of activity, output or outcome indicators (thus excluding routine changes to pay or public funding flows, or user fee regimes).

Another systematic review has recently addressed the use of conditional cash transfers for service users (demand-side paying for performance) for improving the uptake of health interventions in LMICs (Lagarde 2011). This review therefore focuses on evidence of the impacts of supply-side paying for performance aimed at improving the delivery of health interventions. We include all impact evaluations of paying for performance interventions that have a supply-side component.

In this review, paying for performance includes both paying for performance schemes (including ancillary components) and paying for performance per se (where any ancillary components are controlled for).

How the intervention might work

On one level, paying for performance by individuals is not new – it has taken the form of user fees, and in many low- and middle-income countries it remains one of the main forms of health financing. However, public funding for health (including aid funding, where this is channelled through governments) has traditionally not been linked to specific activities, but has taken the form of budget flows, which are linked to indicators such as staffing levels or bed numbers (for facilities), inputs (such as estimated drug needs), population numbers (for regions and districts, in some cases) and also historical trends in expenditure (all modified by overall budget constraints).

These bureaucratic mechanisms offer the advantage of stability and predictability, and rely on local clinical judgement as to how and what services to offer. The disadvantage, however, is that health systems based on budget funding and salaried staff can lack incentives to improve quality, to increase outputs and to improve outcomes. Paying for performance aims to reintroduce those incentives by linking pay (at individual or facility level) to desired activities and/or outcome indicators. It may in addition increase resources (by providing supplementary funding) or may be an alternative mechanism for channelling existing funding resources (substituting for existing funds).

In Organisation for Economic Co-operation and Development (OECD) countries, paying for performance is generally described as a tool for improving quality (Christianson 2007). In LMICs, however, it generally has wider objectives (Eldridge 2009). These include:

- to increase the allocative efficiency of health services (by encouraging the provision of high-priority and cost-effective services);
- to increase the technical efficiency (by making better use of existing resources such as health staff);
- to improve equity of outcomes (for example, by encouraging expansion of services to hard-to-reach groups).

Others emphasise the potential of paying for performance to transform health sectors, introducing client-oriented public finance models inspired by the new public management mode (Meessen 2011).

Paying providers for performance is clearly premised on the assumption that for these three dimensions to shift, a change in behaviour on the provider side is required. If, however, the barriers are more connected with demand-side factors (such as low affordability of services), then paying for performance for providers alone will not be effective.

Paying providers for performance in LMICs can operate at a number of levels. It can be offered directly to health workers (in public, private or private not-for-profit sectors). It can be linked to facility budgets. It can be used to set budgets or supplement budgets at higher organisational units, such as health districts or regions. It can also be used at national level, in particular by donor organisations negotiating aid to a national health sector. Clearly, incentives would be expected to operate differently at these different levels: incentives to individuals are likely to be

more directly motivating (incentives to organisations only affect behaviour indirectly, if passed on in some way to individuals), but may undermine co-operation (unlike organisational incentives, which might be expected to reinforce co-operation).

It seems intuitive that paying more money for the delivery of effective services will improve health care, but health care does not operate like a classic free market. Human behaviour is complex and there are many theories that attempt to explain both health behaviour and professional behaviour. Principal–agent theory addresses relationships where one individual (the patient) cannot directly observe or know the level of skill or effort expended by the other individual (the professional) doing the contracted work. Patients do not have perfect knowledge of their medical condition, their need for care, or the expected outcome of healthcare services, therefore they are willing to have healthcare professionals act as their agents in providing information and services. Patients have asymmetric information about the need for and outcomes of health care, therefore patient demand for health care may be unresponsive to technical quality. Therefore, one theoretical advantage of performance pay is that explicit financial incentives are provided even when patient demand for health care is unresponsive to quality. Professional effort in providing high quality is rewarded, regardless of whether patients recognise it. This theoretical advantage relies, however, on a host of assumptions, including the ability to assess quality, the linkage of paying for performance systems with quality measures, and the absence of adverse consequences.

Moreover, as indicated above, in LMICs in particular, paying for performance is being deployed for a wide range of reasons other than quality. It is envisaged more ambitiously as a tool to increase the responsiveness of staff and the health system generally to priority areas.

It is also important to note that although financial incentives and healthcare payment systems are likely to have an important influence on professional behaviour, this influence is far from exclusive. In economic terms, professionals are viewed as maximising their utility function (i.e. their well-being). Important factors included in their utility function, besides income, include professional and social status (or self image), altruism (doing what they perceive to be best for their patients), the burden of efforts to change their behaviour, and their uncertainty about the benefits of changing their behaviour. Moreover there may be other barriers to changing professional behaviour, even when professionals are motivated, including patient factors, lack of time, lack of technical skills, lack of resources and organisational constraints.

It is generally accepted that professionals are motivated by the satisfaction of doing their jobs well (intrinsic motivation). Indeed, it is doubtful whether some valued-but-difficult-to-observe dimensions of quality (such as empathy or listening in the medical encounter) would be provided at all if physicians were solely interested in income. Health professionals have both monetary and non-monetary incentives, all of which affect their performance. It is possible that financial incentives may dilute professionals' intrinsic motivation and this is the subject of widespread debate around public sector motivation in higher-income countries (Marquand 2004). Psychological studies also highlight the risks to intrinsic motivation of extrinsic rewards (Deci 1999). On the other hand, where health workers' pay is low in absolute terms, incentives may be an important channel to improve motivation through increasing their income levels.

The timescale of evaluation is another important consideration. Financial incentives might be effective in the short run for simple and distinct, well-defined behavioural goals, but these are not necessarily sustained in the longer run.

Paying for performance schemes are often accompanied by ancillary features, such as an increase in resources. When paying for performance schemes are compared to no intervention, it may be impossible to disentangle the impact of paying for performance per se from the impact of increased resources and other ancillary components.

Why it is important to do this review

Although both demand and supply-side paying for performance are widely advocated in LMICs, there are currently no systematic reviews of the impacts of supply-side paying for performance in LMICs (Oxman 2008). Another Cochrane review (Giuffrida 1999) assessed evidence on target payments for primary care. However, this had a more limited scope (focusing on target payments alone - payments for reaching a certain level of coverage, which can be defined in absolute terms or relative to a starting point) and only found studies in high-income countries. With the growth in interest in paying for performance in LMICs, it is believed that more rigorous studies have been produced in the last few years which can now warrant a review focused on the LMICs.

While reviews of schemes in high-income countries can help to inform decisions in LMICs, there are several reasons for undertaking a review of the impacts of paying for performance in LMICs specifically. The potential benefits, harms and costs of paying for performance may be greater in LMICs, where there are fewer resources and financial incentives than in high-income countries, weak health systems, inadequate supplies, facilities and human resources, and greater inequities, and where paying for performance schemes are often introduced by donors and includes ancillary components, such as increased resources and technical support.

Paying for performance is a complex intervention with uncertain benefits and potential harms. It may, just to give one example, lead to the concentration of resources in areas where targets are easier to meet (which typically are better served areas), thus increasing inequity of provision. The extent to which benefits attributed to paying for performance in LMICs are attributable to conditionality (versus ancillary components of paying for performance schemes in LMICs, such as increased resources and technical support) is also uncertain. Paying for performance may not be a good use of resources, even when it is effective, due to potentially small effects and high costs. For these reasons a systematic review of evaluations of the impacts of paying for performance is needed to inform decisions about whether and when to use paying for performance, how to design these schemes, and how to monitor and evaluate them in LMICs.

Finally, this is an area of growing interest for funders and for LMIC governments, so a review of evidence on effectiveness is timely.

OBJECTIVES

The objective of the systematic review is to summarise the current evidence for the effects of paying for performance on the provision of health care and health outcomes in low- and middle-income settings.

METHODS

Criteria for considering studies for this review

Types of studies

The review includes:

- Randomised trials
- Non-randomised trials (experimental studies in which people are allocated to different interventions using methods that are not random)
- Controlled before-after (CBA) studies with:
  * at least two clusters in each comparison group;
  * pre and post intervention periods for study and control groups are the same;
  * appropriate choice of the control site, e.g. similar socio-economic characteristics and/or and no major differences in the baseline group.
- Interrupted time series (ITS) studies with at least three measurements before and after introducing the intervention

Well-designed cluster-randomised trials protect against selection bias and are likely to provide the most rigorous estimates of the impacts of paying for performance schemes. However, cluster-randomised trials may not be practical for evaluating some paying for performance schemes (e.g. when there is simultaneous system-wide implementation). Although CBA studies have a high risk of bias, we believe it is important, at least at this time, to include these studies. ITS studies may be problematic due to changes in information systems and the reliability of information systems used in paying for performance schemes in LMICs. However, they potentially have a lower risk of bias than CBA studies.

Other study designs may provide useful information about acceptability, potential effects or explanations for observed effects of paying for performance, but are unlikely to provide useful estimates of the impact of paying for performance on the main outcomes of this review.

Types of participants

Participants in performance-based financing (PBF) schemes include providers of healthcare services (health workers and facilities), sub-national organisations (health administrations, non-governmental organisations or local governments), national governments and combinations of these. All sectors (public, private and private not-for-profit) have been included in the review.

Types of interventions

Paying for performance takes three main forms.

- Conditional cash payment
- Conditional provision of material goods
- Target payments (payments for reaching a certain level of coverage, which can be defined in absolute terms or relative to a starting point)

We have **included** evaluations of paying for performance schemes (including ancillary components), compared to any alternative (including non-conditional financial incentives and different levels of conditional financial incentives). We have included comparisons with alternatives where there may be differences in ancillary components, such as increased resources, as well as differences in paying for performance.

We have **excluded** studies which focus on the following.

- The demand side of health care only (i.e. payments to consumers, not providers)
- Payment to health workers or facilities which are not explicitly linked to changing patterns of performance (e.g. for coming to work; salary increases; routine increases in activity-based payments such as Diagnosis-Related Groups (DRGs) or fees for service; or changes to budget flows which are routine or intended to motivate, but without being conditional on specific activity or output measures)

Types of outcome measures

Primary outcomes

To be included, a study must report **at least one** of the following outcomes:

- Changes in targeted measures of provider performance, such as the delivery or utilisation of healthcare services, or patient outcomes
- Unintended effects, including motivating unintended behaviours, distortions (ignoring important tasks that are not rewarded with incentives), cherry-picking/cream-skimming (prioritising patients that are most profitable over those who release fewer financial rewards), gaming (improving or cheating on reporting rather than improving performance), increased inequities and dependency on financial incentives
- Changes in resource use, including for incentives, administration and services

Secondary outcomes

The following other outcomes of interest are included if reported in included studies:

- Acceptability
- Patient or provider satisfaction
- Impacts on management or information systems (if not a targeted measure of performance)
- Impacts on overall financing or resource allocation

The results of process evaluations or qualitative studies conducted alongside impact evaluations have been included.

Search methods for identification of studies

Electronic searches

We searched the Database of Abstracts of Reviews of Effectiveness (DARE) for related reviews.

We searched the following electronic databases for primary studies:

- The Cochrane Central Register of Controlled Trials (CENTRAL) 2009, Issue 1, part of the *The Cochrane Library* (www.thecochranelibrary.com) including the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register (searched 3 March 2009)
- MEDLINE, Ovid In-Process & Other Non-Indexed Citations and MEDLINE, Ovid (1948 to present) (searched 24 June 2011)
- EMBASE, Ovid (1980 to 2009 Week 09) (searched 2 March 2009)
- PsycINFO, Ovid (1806 to February Week 4 2009) (searched 4 March 2009)
- EconLit, Ovid (1969 to February 2009) (searched 5 March 2009)
- Sociological Abstracts, CSA (1952 to current) (searched 8 March 2009)
- Social Services Abstracts, CSA (1979 to current) (searched 8 March 2009)
- LILACS (searched 6 May 2009)
- WHOLIS (searched 7 May 2009)
- World Bank
- Science Citation Index Expanded (SCI-EXPANDED) (1975 to present) (searched 8 September 2010)
- Social Sciences Citation Index (SSCI) (1975 to present) (searched 8 September 2010).

In addition we selected relevant databases from the LMIC database list at: http://epocoslo.cochrane.org. We did not search CINAHL or International Pharmaceutical Abstracts, so it is possible that studies relating to nursing or pharmaceuticals were missed. However, the general searches, including in websites focused on this topic, did not suggest that we had missed any relevant studies. We will add these databases when the review is updated.

We developed strategies that incorporated the methodological component of the EPOC search strategy combined with selected index terms and free-text terms. We placed no language or date restrictions on the search strategy. We translated the MEDLINE search strategy into the other databases using the appropriate controlled vocabulary.

The full search strategies for all databases are included in Appendix 1.

Searching other resources

We contacted international experts in the field, including the authors of relevant articles that were retrieved. We asked them to identify additional websites, academic (or other) institutions active in this field and other experts in the field, as well as additional relevant studies.

In addition, we searched the websites of organisations likely to be active in the field, including: the World Bank, the African Development Bank, U.S. Agency for International Development (USAID), Management Sciences for Health (MSH), Centre for Global Development, World Health Organization (WHO), Swiss Tropical Institute, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), KfW Entwicklungsbank, Department for International Development (DFID), The Global Alliance for Vaccines and Immunization (GAVI), The Global Fund to Fight AIDS, Tuberculosis and Malaria, Asian Development Bank and Pan American Health Organization (PAHO).

We also searched the websites of academic institutions active in this field, such as the London School of Hygiene and Tropical Medicine, the Harvard School of Public Health, University of Cape Town, Institute of Policy Studies of Sri Lanka (IPS), the Kenya Institute of Policy Analysis and Research (IPAR), and Institute of Tropical Medicine, Belgium.

We checked references from included studies and other relevant articles to identify other relevant studies that met the inclusion criteria.

We searched ISI Web of Science for papers which cite studies included in the review.

To update our search, prior to finalisation in April 2011, we searched the Results-Based Financing for Health website for any recent studies which we had missed, and re-ran the MEDLINE search in June 2011.

Data collection and analysis

Selection of studies

Two authors independently screened abstracts in order to identify studies which met the inclusion criteria. We then retrieved studies which were selected as meeting or possibly meeting the criteria in full and two authors again checked them independently in order to produce a final list of included studies.

Data extraction and management

Two authors independently extracted the following information from the included studies, using a modified version of the Cochrane EPOC Group data collection checklist:

Study design (randomised trial, non-randomised trial, CBA, ITS)

Type of targeted behaviour (clinical prevention services, diagnosis, test ordering etc.)

Study setting (country, urban, rural)

Participants

- Targets for paying for performance scheme
- Description of patient group(s) affected by intervention

Setting

Methods

- Unit of allocation
- Unit of analysis
- Power calculation
- Quality criteria

Intervention

- Magnitude of incentives
- Incentives relative to appropriate measure (e.g. percentage of health workers' wage)
- Are incentives additional to ordinary wage/funding?

- Ancillary components
- Source of funding

Outcomes

- Main outcome measures
- Economic variables (e.g. change in resource use)
- Length of time during which outcomes were measured after initiation of intervention
- Length of post intervention follow-up period
- Measurement of outcome indicator - by whom?

Results

- Changes in targeted measures of provider performance
- Changes in the delivery of healthcare services
- Changes in the utilisation of healthcare services
- Changes in patient outcomes
- Unintended effects
- Changes in resource use
- Acceptability
- Patient or provider satisfaction
- Impacts on management or information systems (if not a targeted measure of performance)
- Impacts on overall financing or resource allocation
- Equity considerations: evidence of differential impact on different parts of the population

We entered and managed data using Excel.

Assessment of risk of bias in included studies

We used criteria recommended by the Cochrane EPOC Group to assess the risk of bias for each main outcome in all studies included in the review (EPOC Review Group Checklist, 2008).

We assigned an overall assessment of the risk of bias (high, moderate or low risk of bias) to each main outcome in all included studies using the approach suggested in Chapter 12 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2008).

We assessed the quality of evidence for each main outcome – that is the extent of confidence in the estimate of effect across studies - using the GRADE approach (Guyatt 2008).

Measures of treatment effect

For randomised and non-randomised trials and CBA studies we recorded outcomes in each comparison group. Where possible we have recorded or calculated risk ratios (RRs) for dichotomous outcomes. If adjusted analyses are reported (adjusting for potential confounders in non-randomised studies), we used estimates of effect from the primary analysis reported by the investigators and converted these to RRs, if possible.

For ITS studies we recorded changes in level and in slope. If papers with ITS design did not provide an appropriate analysis or reporting of results, but presented the data points in a graph or in a table that could be scanned, we re-analysed the data using methods described in Ramsay 2003. The following segmented time series regression model was specified: Y(t) = B0 + B1*Preslope + B2*Postslope + 3*intervention + e(t), where Y(t) is the outcome in time period t. Pre-slope is a continuous variable indicating time from the start of the study (coded as 0, 1, 2, 3 etc.). Post slope is coded 0 up to and including the first point post intervention and coded sequentially from 1 thereafter. Intervention is coded 0 for pre-intervention time points and 1 for post intervention time points. In this model, B1 estimates the slope of the pre-intervention data, B2 estimates the change in slope post intervention and B3 estimates the change in level of outcome as the difference between the estimated first point post intervention and the extrapolated first point post intervention, if the pre-intervention line was continued into the post intervention phase. The error term e(t) was assumed to be first order autoregressive. For controlled ITS studies, the model also adjusts for the baseline differences between intervention and control sites. If data have been collected on a quarterly basis we have also adjusted for seasonality in the model.

If papers with CBA design did not provide an appropriate analysis or reporting of results, but presented the data for each district/region in the intervention and control groups respectively, we re-analysed the data using a generalised linear model to calculate adjusted RR (Proc GENMOD in SAS version 9.2, SAS Institute Inc., Cary, NC, USA). We created a dataset with the same number of events and non-events per district/region pre- and post intervention as reported in the paper. We estimated the post intervention RR for the event (intervention relative to control), adjusted for the difference in risk between intervention and control pre-intervention, and pre- versus post intervention (underlying trend).

Unit of analysis issues

For cluster-randomised trials and CBA studies we controlled that an appropriate analysis had been done that adjusted for clustering in calculating confidence intervals or P values. If an analysis had not done this, we attempted to extract the necessary data (intra-cluster correlation coefficients - ICCs) or obtain these data from the investigators and reanalyse the results. If this was not possible, we report point estimates, but not the reported confidence intervals or P values.

Dealing with missing data

We have contacted the authors of each of the included studies to obtain missing data, particularly focusing on the results of interventions.

Assessment of heterogeneity

Given the variation found across studies in relation to the intervention design, study design and outcome measures, we have not conducted meta-analysis of results, and a statistical assessment of heterogeneity of results was therefore not done.

Assessment of reporting biases

Selective outcome reporting is a risk for pay for performance studies, where information on many indicators was recorded in the intervention, but not all indicators are reported in studies. Original protocols were not available, despite requesting them from authors in some cases. We assessed risks qualitatively and noted concerns in the review, where a subset of outcomes appeared to be reported.

We also assessed publication bias qualitatively, based on the results and characteristics of the included studies, including the extent to which only small effects in favour of the intervention are reported, the extent to which funders or investigators are advocates

of paying for performance or have a vested interest in the results, and the extent to which the authors' interpretations of the results are supported by the actual results.

Data synthesis

Of the nine studies which were included in the review, only one did not involve important differences in ancillary components. Interpretation of the estimates of the impact of paying for performance therefore has to bear this in mind. We have highlighted differences in context, intervention design, resourcing and ancillary components in the Discussion.

For each of the comparisons, the results of each relevant included study are summarised in a table that includes the main characteristics of the study, the results in natural units as reported by the investigators and standardised results.

Studies of paying for performance are heterogeneous in relation to context, study design, characteristics of the participants and the interventions, and the outcome measures. We therefore judged it to be uninformative to calculate average effects across studies.

Subgroup analysis and investigation of heterogeneity

If we found more than one study for any of the comparisons listed above, we stated that we would explore the extent to which the magnitude of incentives and/or ancillary components might explain differences in the impacts of paying for performance. Of presumed greatest significance was the magnitude and proportion of other financing sources that the incentives constituted (clearly, the higher the incentive, the more we would expect a response from providers).

To highlight the evidence and evidence gaps, we prepared a table indicating the number and type of studies identified, by level of targeting, size of incentives and presence or absence of ancillary measures.

We were to explore heterogeneity visually by preparing tables, bubble plots and box plots to explore the size of the observed effects in relation to each of these variables. We aimed to consider each potential explanatory factor one at a time by looking for patterns in the distribution of the effects of paying for performance. However, because we found a small number of studies relative to the number of potential explanatory factors, we have investigated and reported potential explanations of heterogeneity cautiously, using the criteria for interpreting subgroup analyses and meta-regressions in the *Cochrane Handbook for Systematic Reviews of Interventions* (Section 9.6.6) (Higgins 2008).

Sensitivity analysis

We planned to perform sensitivity analyses by excluding studies with a high risk of bias for any outcome for which we found more than one comparable study with studies with a low or moderate risk of bias. However, due to the assessment of risk of bias being uncertain or high in most of the studies, this was not possible in the end.

RESULTS

Description of studies

Results of the search

Screening of more than 15 databases and the other search strategies yielded 1374 references, which two researchers examined independently. We retrieved 146 complete articles from databases and websites and two authors read these independently. We included 13 studies in the first stage, but at data extraction this was reduced to nine.

Included studies

Nine studies were included (see Characteristics of included studies): one randomised trial (Peabody 2010; Peabody 2011; Valera 2009), six CBAs (Basinga 2010; Canavan 2008; Soeters 2008; Soeters 2009; Soeters 2005; Vergeer 2008) and two ITSs (Liu 2003; Quy 2003). Quy 2003 was not presented as an ITS study, but we have obtained the original data to analyse it in this way.

Intervention characteristics

The Characteristics of included studies tables give a summary of the interventions covered by the nine sets of studies. It is clear that the interventions are varied. One used target payments linked to quality of care (Peabody 2010, alongside an increased access arm). Two used target payments linked to coverage indicators (Canavan 2008; Vergeer 2008). Three used conditional cash transfers, modified by quality measurements (Basinga 2010; Soeters 2009 and 2010). Two used conditional cash transfers without quality measures (Quy 2003; Soeters 2005). One used a mix of conditional cash transfers and target payments (Liu 2003).

The targeted outcomes are also varied. Most of the interventions used a wide range of targets covering inpatient, outpatient and preventive care, including a strong emphasis on services for women and children, paying for outputs or coverage. The exceptions are the Vietnam study, which focused specifically on tuberculosis (cases detected was the main outcome measure); the China study, where no clinical group is specified, since this is a retrospective analysis of hospital bonus payments, using hospital revenue as its main outcome; and the Philippines study, which focused on improved treatment of common illnesses in under-sixes.

Participants and context

A range of participants were targeted by the interventions (Table 1). Four (Canavan 2008; Soeters 2005; Soeters 2008 ; Vergeer 2008) involved a mixed range of public and faith-based facilities (dispensaries, health posts, health centres and hospitals). Two focused on primary facilities alone (Basinga 2010; Soeters 2009). Two focused exclusively on hospitals (Liu 2003; Peabody 2010) and one (Quy 2003) on individual private practitioners.

Although the focus of assessment was the facility or provider level, in the majority of the interventions the programme was managed at a district level (except for the Philippines, China and Vietnam studies). In Tanzania and Zambia, the PBF activities were managed by missionary structures.

In terms of the context, most interventions took place in mostly rural or mixed areas (only one focused on an urban area – the Vietnam intervention focused purely on Ho Chi Minh City).

Target setting and incentive payments

Assessment of performance was carried out by an independent assessor in only one intervention (the Philippines one, which was also the only one designed specifically for study purposes). In the other cases, assessment was by the facilities or providers themselves (Vietnam, China, Zambia and Tanzania); by facilities, supervised by the district (Burundi); and by facilities with periodic verification by external groups (Rwanda, Democratic Republic of Congo).

In relation to the recipients of the incentive payments, these varied (Table 2). The most common arrangement was for incentives to be paid to facilities. In three cases, there was discretion about the amounts passed on to staff by facility managers (Rwanda, Burundi, Democratic Republic of Congo). In two cases, the proportion allocated to staff from within payments to facilities was specified by the donor (Tanzania and Zambia). In three cases, payments went straight to health workers, with a 50:50 split of payments between doctors and other staff in the Philippines and full payments to doctors in China and Vietnam. In three cases, districts were also beneficiaries of incentive payments to cover management costs, though in this case the link to any performance indicators was not clear (Tanzania, Zambia, Democratic Republic of Congo).

The magnitude of incentive payments are not consistently reported, but from the information available it is clear that it varied considerably, ranging from around 5% of doctors' incomes in the Philippines to 60% of total facility income in Burundi.

There is limited information in the studies on how targets were set and also how values were attached to them (where payments are made per act, for example). However, questions are raised by the study authors about the target payments based on coverage in Tanzania and Zambia, including about the accuracy of the population denominators, the fact that coverage targets do not reflect the varied starting point of different facilities, the lack of inclusion of the facility management in setting targets, and the lack of emphasis on preventive care and quality of care measurements.

Ancillary components

All of the interventions had ancillary components, apart from the China study, which was not an intervention as such, but rather a retrospective analysis of changes in bonus payments made to doctors over a period of 22 years. For the other interventions, the most common ancillary components were training, supervision and feedback, although in one case funding for infrastructure was also provided (Zambia). In some cases, such as Zambia and Tanzania, the PBF component was part of a longer-term investment by the donor in the area. Only in the Philippines and Rwanda (Basinga 2010) cases were ancillary components provided equally across intervention and control areas.

In all cases, incentive payments provided additional resources. In only one case (Basinga 2010) was this balanced by equivalent funds provided to the control areas.

Costs and funding

Five out of nine studies gave information on the total cost of the intervention. This links to the funding source, as the five studies with cost information were all funded by the Dutch non-governmental organisation Cordaid, which had set per capita budgets. The range of investment was from USD 0.5 per capita in Tanzania and Zambia to USD 2.6 per capita in Burundi.

Excluded studies

A list of excluded studies is given in the Characteristics of excluded studies table. The most common reason for exclusion was the design of the study.

Risk of bias in included studies

The risk of bias for each of the included studies can be found in the 'Risk of bias' tables. Most interventions have been written up in a number of studies. We cite the main one from which we have drawn evidence on primary outcomes, but have also given details of other studies from which wider data are drawn. For the purposes of assessing risk of bias, Basinga 2010 has been considered a CBA study due to changes which had to be made to the allocation of districts. Sixteen districts were paired, with one in each pair to be randomly allocated to the intervention group and the other to the control. However, some control districts were found to have existing pay for performance schemes and so the allocation had to be adjusted in a non-random way. There are discrepancies in the explanations about how districts were allocated to intervention or control in different reports of the study (Basinga 2009; Basinga 2010; Basinga 2011) and we have not received responses from the investigators to clarify these discrepancies. Consequently, we do not know whether few or many of the random allocations were changed, and whether changes only occurred in the control to intervention direction.

Of the nine studies, we judged the risk of bias as high in seven. Only Peabody 2010 was judged as being at low risk of bias. Allocation was not blinded but we judged this to be unlikely to have biased results. We assessed the risk of bias in Basinga 2010 as unclear, largely due to the changes to allocation of districts. The remaining studies were assessed as high risk. For the CBAs, common concerns include lack of baseline data, possible selective or incomplete outcome reporting, control areas having different characteristics, unclear contamination and concerns about the reliability of chosen outcome indicators. For the two ITS studies, the main concern is that it is unclear whether the intervention was independent of other changes occurring at the same time as the intervention.

Some of these concerns are linked to the nature of the intervention – it is not easy to randomise participants in a pay for performance scheme, nor is it possible to conceal allocation. (Indeed, that would be somewhat contradictory with the intervention, whose whole aim is to motivate behaviour change through payment.) Moreover, these schemes have often been introduced into complex, dynamic environments in which separating the effects of the intervention from other reforms and environmental changes is not easy.

Effects of interventions

See: **Summary of findings for the main comparison**

Reporting of outcomes

Table 3 indicates the outcome measures which studies report on and how these were assessed. Some gaps should be noted. Only one study (Peabody 2010) included any patient health indicators. These were: (1) age-adjusted wasting (as defined in the 2006 World Health Organization Standard Reference), (2) C-reactive protein (CRP) in blood, a possible measure of acute infection, (3) anaemia,

defined as less than 10 g/dl blood haemoglobin, and (4) general self reported health (GSRH), all for under-fives discharged after treatment for diarrhoea or pneumonia. Most studies focused on utilisation or coverage changes (though one study – Liu 2003 – focused on revenues alone). Five studies included some measure of changes to resources (generally changes in out of pocket payments by patients, changes in facility revenues or both). None reported on changes to organisation or delivery of services, on impacts for management and information systems, or on wider impacts for financing or resource allocation, which is surprising given the nature of the intervention.

Unintended effects are also under-reported – only two of the studies (Canavan 2008; Vergeer 2008) considered or looked for any wider effects of the intervention, other than on measures which were targeted by the intervention.

Primary (targeted) outcomes

Effect on utilisation

Seven studies reported effects in terms of changed utilisation of various services. Peabody 2010 reported no increase in the PBF arm, which is as expected given that the intervention is focused on quality improvements rather than utilisation. Basinga 2010 reported the following (see Table 4):

- No impact of PBF on the probability of any prenatal care or on the probability of completing four or more visits.
- A statistically significant impact on the probability of institutional delivery (7% absolute increase, rising from 35% before to 42% after).
- For children, a significant increase in the likelihood of a preventive visit in the four weeks prior to the survey, but no impact on the likelihood of full vaccination.

Results from other studies have to be interpreted with particular caution, given the risk of bias issues highlighted above. Broadly, the Tanzania (Canavan 2008) and Zambia (Vergeer 2008) studies found that performance of the intervention (mission) facilities was similar to or worse than the 'control' government ones for the indicators tracked (outpatients visits, antenatal care, voluntary counselling and testing, inpatients and institutional deliveries).

Based on data obtained from the authors, we present the results from Tanzania in the annexes. Results are presented in natural units and without P values as there were inaccuracies with denominator populations, and could be more than one episode per person in the case of a number of the indicators. Some other data concerns are highlighted below (Table 5; Table 6).

For Vergeer 2008, original data sheets were missing, so the utilisation responses that can be reported are very limited – natural units, or simple directions of change based on visual inspection of figures are reported in the annexes (Table 7). Again, the overall conclusion is clear – there is no clear evidence of improved performance for any of the outcomes in intervention facilities.

Outcomes for which relative risks (RRs) could be calculated, based on the original data, are presented in Table 8. In Burundi, a statistically significant difference was found for institutional deliveries, favouring the intervention sites (RR 1.79), but in the Democratic Republic of Congo, the reverse was found (statistically significant difference, but RR of 0.75). Coverage of bed nets was statistically significantly higher in Burundi (RR 1.9). In Tanzania, inpatient admissions were significantly lower in intervention sites (RR 0.82). In Burundi, pregnant women were statistically significantly more likely to be fully vaccinated (RR 1.13). For all other indicators, no statistically significant difference was found.

The only utilisation indicators which are assessed by more than one study are antenatal care and institutional deliveries (Basinga 2010; Canavan 2008; Soeters 2009; Soeters 2008; Vergeer 2008). Given the different study designs, interventions (including the extent to which control areas were comparable) and data sources for the studies (some using household surveys and others routine facility data), and concerns about risk of bias, we decided not to conduct meta-analysis.

Quality of care

Only three of the studies provide independently assessed measures of quality of care. For Peabody 2010, quality of care was the main focus of the intervention. Using clinical vignettes as a proxy to assess quality, a 10% improvement was found for both intervention arms, compared to baseline levels (the control arm also improved by 6%). Changes were significant for all groups. Improvements were not found until month 12, but were then sustained until the end of the trial (month 36). The effects are attributed to incentives (in the case of Arm B – the bonus group), systems effects (for Arm A – the access group, where increased insurance enrolment fed through into increased hospital revenues) and dissemination (the control group, which received measurement and feedback, in common with the other two arms). The effects of repeat testing were assessed and found not to be significant (Peabody 2011).

In Basinga 2010, quality of prenatal care is assessed by comparing activities undertaken during prenatal visits with the local clinical practice guideline and by investigating whether a tetanus typhoid vaccination was given during prenatal check-ups. Significant improvements in both measures are reported for the intervention group. In addition to these measures, the payment of incentives was linked to a composite quality measure, based on quarterly direct observation by district supervisors and medical records review. The scores for this broader quality assessment are not reported.

In Quy 2003, the main outcome indicator, which can be considered a quality of care outcome, was detection of sputum smear-positive cases in public private mix (PPM) districts (see Table 4). Reanalysis as a controlled interrupted time series (see Table 9, which reports results both for seasonally unadjusted and seasonally adjusted analyses) found no significant shift in level or slope for the intervention area. The short timeframe should be noted, however – the overall period of study was only two years (five quarters before the intervention; three after).

Effect on patient outcomes

Only one study reported on health outcomes for patients (Peabody 2010) (see Table 10). Of the four outcome measures, two showed significant improvement for the intervention group (wasting and self reported health by parents of the under-fives), while two showed no significant difference (being CRP-negative and not anaemic).

Unintended effects

Only two studies reported on unintended effects – in both Zambia and Tanzania there was a concern about the curative nature of the coverage targets and whether this may squeeze out preventive care. However, no conclusive evidence was found to support or refute this concern (Canavan 2008; Vergeer 2008).

Effect on resource use

Most studies reported on changes to resource use (all except Quy 2003). This is unsurprising, given that PBF constitutes a source of additional resources, which would be expected to have knock-on effects for staff, quality of care (e.g. availability of inputs such as drugs) and patients.

In Liu 2003, the authors found that the introduction of bonuses and the switch from bonuses with weaker incentives (e.g. flat rate) to ones with stronger incentives (e.g. related to the quantity of services delivered or the revenues generated by a doctor) were associated with an increased growth in hospital revenues. Our reanalysis using ITS finds that the only group where there is a significant shift in level and slope is group A – the switch from no bonus to flat rate bonuses (see Table 11). To set the estimates in context, the average service revenue in the year before the change in bonus for Group A was CNY 711,725 (Yuan) (real terms, CNY 1975).

The difference may be partly due to the different analysis techniques. It should also be noted that ITS does not allow effectively for exogenous changes, and given that these bonuses were being introduced at different periods over 22 years, it may well be that the results are influenced by wider changes in Shendong Province and China in general. The sequencing of reforms may also play a role – the impact of group A may in part reflect the fact that this was the first wave of reforms to bonuses, taking place in the early to mid-1980s, while other bonus types were typically introduced later.

The outcome measure in this case – hospital revenues – is in itself an ambiguous indicator. Higher revenues might be good for the hospital but they are likely to indicate problems of affordability and lack of financial protection for patients.

Given the large number of indicators tracked overall in these PBF interventions, we have focused our efforts for other studies on reanalysing patient and utilisation outcomes. For resources, we draw on the original reports. We have reanalysed resource use for Liu 2003, however, as this was its main and only outcome indicator. Changes to resource use in other studies are summarised in Table 4. Lack of baseline data is a challenge for some studies (Canavan 2008; Soeters 2005; Vergeer 2008). In others, results are mixed. For example, in the Democratic Republic of Congo, payments by patients were reported to increase in the intervention group (Soeters 2008), while in Burundi, they were reported to decrease (Soeters 2009). These differences are most likely explained by the differences in starting levels and also the heterogeneity not only of the intervention but also the support which was provided to the 'control' areas. In all studies which monitored overall health facility revenues (Soeters 2005; Soeters 2008; Soeters 2009), the PBF intervention was associated with increases in external support, with knock-on effects for staffing and staff pay, compared to control areas. In Rwanda (Basinga 2010), the amounts paid through PBF were matched for control facilities, but other resource variables (total facility revenues; patient payments) were not reported.

In the Philippines, the relationship between quality scores at facility level and patient costs was investigated. A u-shaped relationship was found, with costs falling as scores rose from below average to average, but then increasing as the quality scores rose above the average mark. At low levels of quality (below 60%), every 10% point increase in quality was associated with an average 20% decline in charges; thereafter every 10% point increase was associated with an average 22% increase in charges (Peabody 2010). This is promising, in that it indicates potential gains to costs and efficiency from quality enhancement measures (which in this study were equally achieved through PBF and increased access/insurance arms).

Secondary outcomes

A summary of findings on secondary outcomes is presented in Table 12.

Patient and provider satisfaction

Five studies reported on patient satisfaction (or assessments of quality of care) and feedback from managers and health workers. The view from patients is mixed – improvements reported in the bonus group in the Philippines (Peabody 2010); a deterioration in perceived quality of care for both groups in Burundi; no difference in perceptions between PBF and non-PBF in Tanzania; concerns expressed by both groups in Zambia; and improvements in perceived quality of care by patients in PBF areas in the Democratic Republic of Congo.

The staff perspective is reported in four studies. In Tanzania, satisfaction was similar for both groups (staff in PBF and non-PBF facilities) and staff expressed concerns about how the PBF targets were set. In Zambia, health workers expressed concerns about workload and the differential bonuses generated by the PBF scheme. In Rwanda (2005), many staff comments related to the particular NGO and its mode of support, but in general there was satisfaction with PBF (though questions about specific payments per indicator). In the Democratic Republic of Congo, managers in both groups expressed dissatisfaction.

Equity

Only one study looked specifically at equity, by disaggregating changes to household payments in the PBF areas (Soeters 2008). Household payments as a proportion of income by the poorest were found to reduce more in intervention districts (by 63.5%, compared to a 21.9% reduction in controls), though it should be noted that payments in PBF areas were at much higher level to start with. For the poor (second quartile), the reverse was found, with a 76.5% reduction in the controls, compared to a 36.2% reduction in intervention district households.

DISCUSSION

The impact of the performance-based financing (PBF) is likely to depend on a range of factors, including how it is designed, the degree of participation in setting targets, what targets are used, how they are measured, the level of rewards they attract and by the context in which the PBF takes place, including the efficiency of implementation systems and underlying factors such as starting levels of pay and funding. For that reason, this review has presented

considerable detail on the implementation of the PBF schemes, as these factors are key to interpreting the results.

The fundamental idea behind PBF is to align the incentives of providers with those of the commissioner of care. However, the mechanisms through which the interventions might work were not always clear. In Peabody 2010a, for example, payments to doctors were linked to proxy quality assessments. Every six months, a small group of doctors per hospital were randomly tested using clinical practice vignettes, which test clinical knowledge. If they met the thresholds, the bonuses were paid to all doctors in that hospital. No training was provided by the programme, so the channels by which quality might have risen were not clear. Presumably the testing system may have motivated collective actions to raise standards, such as training. It was striking, however, that the direct incentives arm (PBF) achieved identical improvements in quality to the indirect incentives arm (increased access).

One of the review expectations was that the scale of incentives would be the main factor affecting impact, with more substantial incentives more likely to motivate behaviour change. There are insufficient robust studies to make a fair comparison, but the only study assessed as having a low risk of bias (Peabody 2010) achieved statistically significant gains in two out of four outcome measures, despite the small magnitude of payments (an estimated 5% of doctors' income).

By contrast, the 2010 Rwanda study (Basinga 2010) concluded that the reason why some targeted indicators showed significant gains and others not was that the latter were too poorly remunerated (relative to effort they required), were influenced by client rather than provider actions, or both. This apparent difference in finding may relate to the different targets (paying for quality versus paying per output) and other factors, such as different starting points for salaries, which are not reported, and the different service types (largely preventive in Rwanda; child curative care in the Philippines). One author, for example, recommends that variable incentives should be considerable and not substantially smaller than the minimum salaries that are usually fixed by government (Soeters 2008). This presumably reflects the low salary levels in that context (the Democratic Republic of Congo).

The question of magnitude of payments raises a boundary issue which was challenging for the review – is PBF just another form of provider payment, like Diagnostic Related Groups or contracting, in which the full cost of service is transferred, but linked to specified outputs? It is sometimes used in this way, for example in relation to performance-based contracts. However, others use it to describe marginal payments that do not replace regular funding systems (or cover full costs of service provision) but aim to modify provider behaviour. The studies included here all fall into the second category, even though the magnitude of transfers varied considerably. The Democratic Republic of Congo study makes the point that subsidies do not require cost information, but can reflect the services which are lagging relative to national targets, and the availability of funds (Soeters 2011).

Proponents of the PBF approach claim that it focuses on outputs/outcomes and allows managers free rein to manage. However, micro-monitoring is often involved in the studies included in this review, which is not necessarily distinguishable from micro-management. To give an example from Burundi, in addition to specified payments per case, a bonus of up to 15% of total payments was offered for quality. This was assessed quarterly, using a composite quality index with 153 indicators. One of the indicators is "Incinerator and placenta pit within fence and locked". To obtain the points allocated for this indicator, all these criteria must be met: (a) functional cleaned incinerator; (b) placenta pit available with lid and; (c) a well-built fence with door closed under key (Soeters 2010). This level of detail is clearly focusing not just on outputs or outcomes, but on internal and organisational processes. There is little evidence from these studies that PBF has acted as a trigger for increased organisational or managerial autonomy.

There is some suggestion that PBF approaches have been more effective in post conflict areas (e.g. Rwanda and Cambodia) (Toonen 2009). If this is indeed the case, it may be because new structures and payment systems are easier to establish in health systems which are in flux. It is also possible that this pattern reflects the access and role of donors and international non-governmental organisations (NGOs) in areas where government systems are less established.

Summary of main results

Overall, the quality of evidence is graded as low or very low (see Summary of findings for the main comparison), with limited numbers of studies reporting on specific indicators, high risk of bias in most studies, and inconsistency of findings. There is some evidence that some schemes have had some success, but findings vary in direction of change as well as magnitude across settings, and so it is not possible to draw any general conclusions. Although it seems likely that the impacts of paying for performance will depend on the design of the intervention, ancillary components of pay for performance schemes, and contextual factors, the studies included in this review do not provide a solid basis for guiding the design of pay for performance schemes or determining the contexts in which they are more or less likely to be effective or to have adverse effects.

We conclude that there are few robust studies of PBF available from a low- or middle-income context and it is premature to draw any firm conclusions on its effectiveness or factors that determine its effectiveness. This is, however, an area receiving considerable interest, both in terms of implementation and research, and more studies are expected over the next few years. There is considerable scope for improvements in study designs, including in relation to making a broader assessment of unintended consequences, and differentiating the effects of ancillary components from those of PBF incentives.

Overall completeness and applicability of evidence

The studies included here were limited in number, quality and range of schemes and contexts covered. The evidence base is therefore incomplete and any findings to date have limited applicability elsewhere. In particular, it is clear that studies of interventions targeting the sub-national and national level are absent (see Table 13). This largely reflects study design issues: output-based aid approaches have not been studied using randomised or non-randomised trial, controlled before-after (CBA) or interrupted time series (ITS) methodologies.

In addition, as with other fields (Hopewell 2009), there may be a bias against publishing "negative" findings. For example, an unpublished World Bank report of a quasi-experimental study of a PBF mechanism started in 2003 in Uganda was available in

2007 (Morgan 2010), but has not been published. This might be related to the findings of the study, which is reported to have "found no discernable impact of bonuses on the provision of health services" (Morgan 2010).

Subgroup analysis

Table 13 gives a summary of studies which were included and how they fit the analysis grid defined in the protocol.

In relation to the level at which incentives were targeted, this is not entirely clear in all cases, as in a number of interventions, the primary level receiving the payments was the facility, but with some trickle-down to staff (predefined or not) and with some additional payments being made to the districts in three cases. However, since the unit of performance was the facility, these interventions have been classified as targeting the facility level. This is the dominant mode for included studies (six out of nine targeting the facility level, and three the individual).

In relation to ancillary components, most interventions included ancillary interventions which are not adjusted for in relation to the control groups (or the ancillary components are not known, as in Liu 2003, which is a retrospective study with very little contextual information). This makes it hard to judge the additional effects of the PBF component. In the one study which did adjust, the magnitude of incentives was substantial (Basinga 2010).

Quality of the evidence

The overall quality of current evidence is poor (see Table 14 for a summary of study characteristics and comments on study designs). Only one study was assessed as having low risk of bias.

Problems which were common amongst the studies identified included the following.

- Non-random allocation of the intervention
- Additional resources and ancillary components that may be responsible for impacts rather than conditional payments
- Other confounders (e.g. contextual differences between intervention and non-intervention groups)
- Lack of rigorous before and after measures of effect
- Lack of consideration of wider systemic issues (e.g. adverse impacts on other services)
- A plethora of targets (outcome measures) and consequently a high risk of selective outcome reporting (i.e. reporting statistically significant results and not reporting results that are not statistically significant)
- Conflicting interests (due to PBF being evaluated by individuals and organisations that are advocating and implementing it)
- Not measuring health outcomes, which means that the relationship between process measures and health benefits in many cases is uncertain

Potential biases in the review process

No potential biases in the review process have been identified. However, some improvements to the search strategy are noted for the revised review. In order to identify any studies focusing on nurses or pharmacists, the next review should add these as search terms, and the CINAHL and International Pharmaceutical Abstracts should be added to the databases for searching. PBF as a text word search should also be included, as this is becoming the most common acronym used in the field.

Agreements and disagreements with other studies or reviews

There is a burgeoning literature on PBF, much of it focused on the theory behind PBF (Savedoff 2010), on case studies and on implementation issues. There is an enthusiasm to explore the potential of PBF, including to use it as leverage for wider systems reforms (Meessen 2011). There is, however, as yet limited robust evidence of impact in low- and middle-income countries. This review is therefore contributing to a growing field. Recent studies reviewing the evidence concur that the current evidence base is weak (Sheffler 2010).

Ongoing studies

There are a number of ongoing studies of which we are aware. The World Bank is supporting a number of PBF impact evaluations - for example, one is ongoing in Haut Katanga, Democratic Republic of Congo. Cordaid is also supporting impact evaluations in a number of locations, including the Central African Republic. The Norwegian government is supporting impact evaluations in several countries, including Ghana and Tanzania.

AUTHORS' CONCLUSIONS

Implications for practice

The current evidence base for performance-based financing (PBF) is weak, so any conclusions on implications for practice have to be tentative.

There is some evidence of potential for improving health outcomes, although this is based on one study of one set of services, in which direct and indirect incentives appeared to achieve similar results and in which the mechanisms for achieving change were not clear. In this study, gains were found for self reported general health of children and wasting, but not for anaemia and being C-reactive protein (CRP)-negative. In another study with a moderate risk of bias, significant gains were achieved in some indicators (four out of seven assessed), focused on utilisation of maternal and child health services, but including one 'quality of care' indicator (likelihood of receiving tetanus typhoid vaccination during antenatal care). Other studies found mixed results and presented a range of difficulties in relation to their design. Longer-term questions of sustainability and cost-effectiveness are untested, and adverse consequences have not been considered in most cases.

Implications for research

There is a need for high-quality research into PBF, looking at a range of modalities, scales of magnitude, levels of implementation within the health system, types of services and providers, and contexts. More robust study designs are needed to detect reliably the small to moderate, though potentially important, effects that PBF may be expected to have. Although there may be practical challenges with implementing randomised experiments of PBF, recent studies have demonstrated that it can be feasible to conduct such trials. Robust effectiveness evaluations should be complemented by in-depth process evaluations to uncover the mechanisms by which the intervention may or may not work, and to probe the motivational effects which are intended to be at the core of the intervention.

It will also be important to study the changing effects of incentives over a longer-term time span. As targets are adjusted over time, for example, will gains (where gains were made) be maintained? This could not be assessed as all studies covered a short time-span and focused on introduction of incentive payments, not their removal or change. The financial sustainability of schemes, especially those which are externally funded, should also be investigated.

Another important area for future research is that of the cost-effectiveness of PBF approaches. For interventions which were non-governmental organisation (NGO) projects, overall cost figures were provided in these studies, but the more robust studies did not specify the additional investment, and how this compares with marginal gains.

The impact of PBF on user costs is discussed in a few of the studies, but this is another area which merits more serious study. Are gains to providers retained, or are they passed on to users in order to attract their business? The equity impact of PBF schemes is another of the many under-researched areas - what are the differential impacts on different social groups?

As noted, evaluations should take a broad perspective and consider wider health systems effects, intended or unintended. No organisational impacts were reported in the included studies, despite the view that PBF can increase managerial autonomy.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of Marit Johansen in developing and implementing the search strategies, of Andy Oxman in supporting the development of the review, of Kjetil Olsen for helping with obtaining and organising the studies, and of Jan Odgard-Jensen for his assistance with the statistical analysis. We also acknowledge the helpfulness of the original study authors in providing additional data and information. Additional data and/or papers were provided by Ann Canavan, Xingzhu Liu, Knut Lonroth, John Peabody, Robert Soeters and Petra Vergeer.

REFERENCES

References to studies included in this review

Basinga 2010 *{published data only}*

Basinga P, Gertler P, Binagwaho A, Soucat A, Sturdy J, Vermeersch C. Paying primary health centres for performance in Rwanda. World Bank, Washington, DC, Policy research working paper 5190.

Canavan 2008 *{unpublished data only}*

Canavan A, Swai G. Payment for Performance (P4P) Evaluation: Tanzania Country Report for Cordaid Godfrey Swai, National Consultant Width 1. KIT, Amsterdam.

Liu 2003 *{published data only}*

Liu X, Mills A. The influence of bonus payments to doctors on hospital revenue: results of a quasi-experimental study. *Applied Health Economics & Health Policy* 2003;**2**:91-8.

Peabody 2010 *{unpublished data only}*

* Peabody J, Shimkhada R, Quimbo S, Solon O, Javier X, McCulloch C. Linkage between measurement and incentive payments and health outcomes: experimental data from a policy experiment in the Philippines. Submitted for publication to JAMA.

Quy 2003 *{published data only}*

Quy H, Lan N, Lonnroth K, Buu T, Dieu T, Hai T. Public-private mix for improved TB control in Ho Chi Minh City, Vietnam: an assessment of its impact on case detection. *International Journal for Tuberculosis and Lung Disease* 2003;**7**:464-71.

Soeters 2005 *{unpublished data only}*

Soeters R, Musango L, Meessen B. Comparison of two output based schemes in Butare and Cyangugu provinces with two control provinces in Rwanda. GBPOA, World Bank, Ministry of Health Rwanda 2005.

Soeters 2008 *{unpublished data only}*

Soeters R, Kimakuka C. Résultats de l'enquête ménage, l'enquête qualité, et l'enquête infirmiers titulaires. Pour le Programme Achat de Performance dans les Zones de Santé du District Sanitaire Nord du Sud Kivu February 2008.

Soeters 2009 *{unpublished data only}*

Soeters R, Kiwanuka C. [Rapport de l'Etude d'Evaluation du programme Achat de Performance dans les Provinces Bubanza et Cankuzo, basé sur les résultats des enquêtes ménages, qualité et infirmiers titulaires réalisées en 2006 et 2008]. Report for Cordaid 2009.

Vergeer 2008 *{unpublished data only}*

Vergeer P, Chansa C. Payment for Performance (P4P) Evaluation: Zambia Country Report for Cordaid. KIT, Amsterdam.

References to studies excluded from this review

Bhushan 2002 *{published data only}*

Bhushan I, Keller S, Schwartz B. Achieving the twin objectives of efficiency and equity: contracting health services in Cambodia. ERD Policy Brief Series: Economics and Research Department - Number 6. Asian Development Bank, 2002.

Biai 2007 *{published data only}*

Biai S, Rodrigues A, Gomes M, Ribeiro I, Sodemann M, Alves F, et al. Reduced in-hospital mortality after improved management of children under 5 years admitted to hospital with malaria: randomised trial. *BMJ* 2007;**335**(7625):862.

Chee 2004 *{published data only}*

Chee G, Fields R, Hsi N, Schott W. Evaluation of GAVI immunization services support funding. The Global Alliance for Vaccines and Immunization 2004.

Chee 2007 *{published data only}*

Chee G, Hsi N, Carlson K, Chankova S, Taylor P. Evaluation of the First Five Years of GAVI Immunization Services Support Funding. Bethesda, MD: ABT Associates Inc September 2007.

Chee 2008 *{published data only}*

Chee G, Molldrem V, Hsi N, Chankova S. Evaluation of the GAVI Phase 1 Performance (2000-2005). Bethesda, MD: ABT Associates Inc October 2008.

Eichler 2007 *{published data only}*

Eichler R, Auxila P, Antoine U, Desmangles B. Performance-Based Incentives for Health: Six Years of Results from Supply-Side Programs in Haiti. Center for Global Development, Working Papers #121 2007.

Furth 2006 *{published data only}*

Furth R. Zambia pilot study of performance-based incentives. USAID 2006.

Gauri 2004 *{published data only}*

Gauri V, Cercone J, Briceno R. Separating financing from provision: evidence from 10 years of partnership with health cooperatives in Costa Rica. *Health Policy and Planning* 2004;**19**(5):292-301.

Keller 2001 *{published data only}*

Keller S, Schwartz JB. Final evaluation report: contracting for health services pilot project (CHSPP) - a component of the basic health services project. ADB Loan No. 1447-CAM. Phnom Penh, Cambodia, 2001.

Lim 2008 *{published data only}*

Lim SS, Stein DB, Charrow A, Murray CJ. Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage. *Lancet* 2008;**372**(1474-547X (Electronic), 0140-6736 (Linking), 9655):2031-46.

Liu 2005 *{published data only}*

Liu X, Mills A. The effect of performance-related pay of hospital doctors on hospital behaviour: a case study from Shandong, China. *Human Resources for Health [Electronic Resource]* 2005;**3**:11.

Lu 2006 *{published data only}*

Lu C, Michaud CM, Gakidou E, Khan K, Murray CJ. Effect of the Global Alliance for Vaccines and Immunisation on diphtheria, tetanus, and pertussis vaccine coverage: an independent assessment. *Lancet* 2006;**368**:1088-95.

Mahmood 2007 *{published data only}*

Mahmood J, Alam MN. Bangladesh: Urban Primary Health Care Project. Asian Development Bank 2007; Vol. 29033.

Marek 1999 *{published data only}*

Marek T, Diallo I, Ndiaye B, Rakotosalama J. Successful contracting of prevention services: fighting malnutrition in Senegal and Madagascar. *Health Policy Plan* 1999;**14**(0268-1080 (Print), 0268-1080 (Linking), 4):382-9.

Meessen 2006 *{published data only}*

Meessen B, Musango L, Kashala JP, Lemlin J. Reviewing institutions of rural health centres: the Performance Initiative in Butare, Rwanda. *Tropical Medicine and International Health* 2006;**11**(8):1303-17.

Meuwissen 2006 *{published data only}*

Meuwissen LE, Gorter AC, Kester ADM, Knottnerus JA. Can a comprehensive voucher programme prompt changes in doctors' knowledge, attitudes and practices related to sexual and reproductive health care for adolescents? A case study from Latin America. *Tropical Medicine and International Health* 2006;**11**(6):889-98.

Schwartz 2004 *{published data only}*

Schwartz JB, Bhushan I. Improving immunization equity through a public-private partnership in Cambodia. *Bulletin of the World Health Organization* 2004;**82**(0042-9686 (Print), 0042-9686 (Linking), 9):661-7.

Soeters 2003 *{published data only}*

Soeters R, Griffiths F. Improving government health services through contract management: a case from Cambodia. *Health Policy and Planning* 2003;**18**(1):74-83.

Sondorp 2009 *{published data only}*

Sondorp E, Palmer N, Strong L, Wali A. Afghanistan: Paying NGOs for performance in a postconflict setting. Center for Global Development 2009.

World Bank 2003 *{published data only}*

World Bank. Partnerships within the public sector to achieve health objectives - Performance-based management in an evolving decentralized public health system in West Africa: the case of Burkina Faso. World Bank 2003.

Yip 2001 *{published data only}*

Yip W, Eggleston K. Provider payment reform in China: the case of hospital reimbursement in Hainan province. *Health Economics* 2001;**10**(4):325-39.

References to studies awaiting assessment

Huntingdon 2010 *{published data only}*

Huntington D, Zaky HH, Shawky S, Fattah FA, El-Hadary E. Impact of a service provider incentive payment scheme on quality of reproductive and child-health services in Egypt. *Journal of Health, Population & Nutrition* 2010;**28**:273-80.

Lundberg 2007 *{unpublished data only}*

Lundberg M, Marek T, Okwero P. Contracting Health Services in Uganda, Unpublished Report. World Bank 2007.

Additional references

Basinga 2009

Basinga P. Impact of performance-based financing on the quantity and quality of maternal health services in Rwanda. PhD Dissertation, School of Public Health and Tropical Medicine, Tulane University.

Basinga 2011

Basinga P, Gertler P, Binagwaho A, Soucat A, Sturdy J, Vermeersch C. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet* 2011;**377**:1421-8.

Christianson 2007

Christianson J, Leatherman S, Sutherland K. Financial incentives, health care providers and quality improvements: a review of the evidence. London: The Health Foundation, 2007.

Deci 1999

Deci E, Ryan R, Koestner R. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 1999;**125**(6):627-68.

Eichler 2006

Eichler R. Can "Pay for Performance" increase utilization by the poor and improve the quality of health services? Discussion paper for the first meeting of the Working Group on Performance-Based Incentives. Washington DC, Center for Global Development 2006.

Eldridge 2009

Eldridge C, Palmer N. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. Health Policy and Planning 2009. In press.

Giuffrida 1999

Giuffrida A, Gosden T, Forland F, Kristiansen I, Sergison M, Leese B, et al. Target payments in primary care: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 1999, Issue 4. [DOI: 10.1002/14651858.CD000531]

Guyatt 2008

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**26**:924-6.

Higgins 2008

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.1 (updated September 2008), The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Hopewell 2009

Hopewell S, Loudon K, Clarke M, Oxman A, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.MR000006.pub3]

Lagarde 2011

Lagarde M, Palmer N. The impact of user fees on access to health services in low- and middle-income countries. *Cochrane Database of Systematic Reviews* 2011, Issue 4. [DOI: 10.1002/14651858.CD009094]

Marquand 2004

Marquand D. Decline of the Public. The Hollowing Out of Citizenship. Polity Press, 2004.

Meessen 2011

Meessen B, Soucat A, Sekabaraga C. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform?. *Bulletin of the World Health Organization* 2011;**89**:156.

Morgan 2010

Morgan L. Some days are better than others: Lessons Learnt from Uganda's First Results Based Financing Pilot, Feature story. Washington DC: The World Bank 2010.

Oxman 2008

Oxman AD, Fretheim A. An overview of research on the effects of results-based financing. Report No. 16-2008. Oslo: Norwegian Knowledge Centre for the Health Services 2008.

Peabody 2010a

Peabody JW, Florentino J, Shimkhada R, Solon O, Quimbo S. Quality variation and its impact on costs and satisfaction: evidence from the QIDS study. *Medical Care* 2010;**48**:25-30.

Peabody 2011

Peabody J, Shimkhada R, Quimbo S, Florentino J, Bacate M, McCulloch C, et al. Financial incentives and measurement improved physicians' quality of care in the Philippines. *Health Affairs* 2011;**30**:773-81.

Ramsay 2003

Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. Interrupted time series design in health technology assessment: Lessons from two systematic reviews of behaviour change strategies.. *International Journal of Technology Assessment in Health Care* 2003;**19**:612–23.

Savedoff 2010

Savedoff W. Basic Economics of Results-based Financing in Health. Bath, Maine (USA): Social Insight, 2010.

Sheffler 2010

Sheffler R. Pay For Performance (P4P) programs in health services: what is the evidence?. World Health Report 2010 Background Paper 31, Geneva, WHO 2010.

Soeters 2010

Soeters R, Basenya O, Kamana J, Nyarushatsi JP, Bossuyt M, Vroeg P. Delivering equitable and quality health services through performance based financing. An experiment - control study from Burundi. Submitted for publication to Health Policy and Planning.

Soeters 2011

Soeters R, Peerenboom P, Mushagalusa P, Kimanuka C. Performance-Based Financing Experiment Improved Health Care In The Democratic Republic Of Congo. *Health Affairs* 2011;**30**:1518-27.

Toonen 2009

Toonen J, Canavan A, Vergeer P, Elovainio R. Performance Based Financing: a synthesis report. KIT, in collaboration with Cordaid and WHO, Amsterdam 2009.

Valera 2009

Valera M, Quimbo S. A quality improvement demonstration study: a P4P experiment in the Philippines. Presentation to Asia Pay for Performance Workshop, Cebu, Philippines 2009.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Basinga 2010

Methods	CBA - Rwanda Districts randomly allocated to PBF group or to control in 2006 but several allocations were changed due to a political decision to implement PBF in some of the study's control areas. Bonuses paid to health centres for quantity and quality outputs. Control received same average amount as PBF districts in addition to previous budgets, per quarter
Participants	Primary care facilities

Basinga 2010 *(Continued)*

Interventions	Conditional cash transfers, modified by quality measurement
Outcomes	14 maternal and child health outputs, including curative visits, preventive visits, ANC, FP, vaccination, deliveries and referrals
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"For all district pairs, a coin was flipped to decide which area would be randomly assigned to treatment, and which to control" (Basinga 2009)
Allocation concealment (selection bias)	Low risk	Not applicable as sequence generation and allocation happened simultaneously
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Not possible and not considered a potentially important source of bias
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Not done, but not likely to be an important source of bias
Incomplete outcome data (attrition bias) All outcomes	Low risk	All clusters remained in trial
Selective reporting (reporting bias)	Low risk	No evidence of selective outcome reporting
Other bias	Unclear risk	Main source may be selection bias due to change in allocation (original random allocation of districts had to be altered, as some districts allocated to control group were already in the process of implementing PBF projects with NGO support)
Baseline measurement	Low risk	Household studies conducted prior to intervention
Matched characteristics for control study sites	Low risk	Characteristics compared; analysis allows for differences
Protection against contamination (intervention and controls)	Unclear risk	No information provided by authors
Reliable primary outcome measure	Low risk	Based on household self reporting
Overall assessment of risk of bias	Unclear risk	Our overall assessment is that it unclear to what extent the results from this study are biased

Canavan 2008

Methods	CBA - Tanzania

Canavan 2008 *(Continued)*

	Intervention facilities paid fixed amount at start of year. Equivalent value is available retrospectively if targets are met. Five targets used, each with equal weight and set at same level for all facilities. Performance compared retrospectively with selection of government facilities.
Participants	Mission hospitals, health centres and dispensaries in 5 dioceses
Interventions	Target payments
Outcomes	1. New VCT clients of 10 per 1000 population at hospital and 20 at HC/dispensaries 2. OPD rates of 0.6 per capita (HCs and dispensaries) 3. IPD of 40 per 1000 (hospitals) 4. Institutional deliveries of 10 per 1000 (hospitals) and 20 per 1000 (HCs and dispensaries) 5. Availability of 10 essential drugs
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not done
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Many indicators are reported; it is not easy to assess their completeness and how denominators have changed over the study period
Selective reporting (reporting bias)	Unclear risk	Due to the large number of outcome measures, we cannot exclude the possibility of selective reporting
Other bias	Low risk	None identified
Baseline measurement	High risk	Not done
Matched characteristics for control study sites	High risk	Sites chosen retrospectively; no evidence of matched characteristics
Protection against contamination (intervention and controls)	Unclear risk	No information provided
Reliable primary outcome measure	High risk	Outcome measure not independently assessed
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Liu 2003

Methods	ITS - China Retrospective study of relationship between bonuses paid by county hospitals in one province of China and hospital revenues. There were 6 different comparison groups, A-F. Group A compared a "flat bonus" ("a certain percentage of the hospital's revenue was distributed almost evenly to all staff member every month if the revenue target of the hospital was reached") with a no bonus group. Group B compared a 'quantity-related bonus' with no bonus. Group C compared a 'revenue-related bonus' with

Liu 2003 *(Continued)*

	no bonus. Group D compared the quantity-related bonus with the flat bonus. Group E compared the revenue-related bonus with the flat bonus. Group F: quantity-related compared to revenue-related bonus
Participants	Doctors
Interventions	Mostly conditional cash transfers, but the flat bonus can be seen as a form of target payment
Outcomes	Hospital revenues
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding of outcome assessment (detection bias) All outcomes	Low risk	No blinding but detection bias is unlikely for this outcome measure
Selective reporting (reporting bias)	Low risk	No evidence of this
Other bias	Low risk	No evidence of this
Reliable primary outcome measure	Low risk	Outcome measure is judged to be reliable
Intervention independent (ITS)?	High risk	Study covers a period of major reform. We cannot rule out effects of other changes
Analysed appropriately (ITS)?	Low risk	Original data reanalysed by review authors
Shape of effect pre-specified (ITS)?	Low risk	Done
Reasons given for number of points pre and post (ITS)	Unclear risk	No reasons given
Unlikely to affect data collection (ITS)?	Low risk	Hospital revenue data are unlikely to have been affected by intervention
Incomplete outcome data addressed (ITS)?	Low risk	Dataset appears to be complete
Overall assessment of risk of bias	Unclear risk	Our overall assessment is that it unclear to what extent the results from this study are biased

Peabody 2010

Methods	RCT - Philippines Intervention designed as a study. Three groups: bonus payments for doctors meeting higher quality of care in group 1; increased enrolment into Phil Health Insurance for indigent children under 6 in group 2; no change group 3

Peabody 2010 *(Continued)*

Participants	Public district hospitals
Interventions	Target payments
Outcomes	Improved management of common childhood illnesses; patient outcomes and patient satisfaction
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Hospitals matched in groups of 3, and then allocated randomly to 1 of the 3 groups
Allocation concealment (selection bias)	Low risk	Not applicable as sequence generation and allocation happened simultaneously
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Participants not blinded but this is not considered likely to introduce bias
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not done
Incomplete outcome data (attrition bias) All outcomes	Low risk	This criterion did not apply as the study did not follow a cohort. The response rate was acceptable
Selective reporting (reporting bias)	Low risk	No evidence of this
Other bias	Low risk	None identified
Baseline measurement	Low risk	Done
Matched characteristics for control study sites	Low risk	Done
Protection against contamination (intervention and controls)	Unclear risk	Measures to protect against contamination are not adequately described
Reliable primary outcome measure	Unclear risk	Some of the outcome measures are self reported and the quality measure is still being validated as a reliable proxy
Overall assessment of risk of bias	Low risk	Our overall assessment was that there was a small risk of biased results from this study

Quy 2003

Methods	Analysed as ITS - Vietnam

Quy 2003 *(Continued)*

	Part of public private mix project, started in 2001. Private practitioners invited to join, which meant that 1) participants requested to adhere to National TB programme guidelines (sputum microscopy for diagnosis, referral to laboratories). Participants received special referral forms and quarterly supervisions visits; and 2) were paid fixed incentive payments. There was no control group as such, but information from other districts in Ho Chi Minh City was used to indicate the trend without the intervention
Participants	Private doctors and pharmacists (latter only able to refer, not diagnose and treat)
Interventions	Conditional cash transfers
Outcomes	Increased testing, treatment and referral of TB patients
Notes	

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Not possible, done but unlikely to introduce bias
Incomplete outcome data (attrition bias) All outcomes	Low risk	Not applicable to this study design
Selective reporting (reporting bias)	Low risk	No evidence of this
Other bias	Low risk	None identified
Reliable primary outcome measure	Unclear risk	Outcome indicator (TB cases referred) may be subject to gaming
Intervention independent (ITS)?	Unclear risk	Other contextual factors cannot be assessed
Analysed appropriately (ITS)?	Low risk	Data reanalysed by review authors
Shape of effect pre-specified (ITS)?	Low risk	Done
Reasons given for number of points pre and post (ITS)	Unclear risk	No reasons given
Unlikely to affect data collection (ITS)?	Low risk	This is judged unlikely
Incomplete outcome data addressed (ITS)?	Low risk	No evidence of gaps in data
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Soeters 2005

Methods	CBA - Rwanda Two provinces implemented PBF, but not in same way (different targets and payments per target in each, and supported by different international NGOs). Study compared 4 provinces: 2 implementing ones, 1 supported by bilateral donor with no PBF, 1 with no donor support, but more government staff
Participants	Public health centres only in Butare, but in Cyangugu also district hospitals
Interventions	Conditional cash transfers
Outcomes	Butare: 5 indicators, including new OPD cases; institutional deliveries; referred deliveries; new FP users; measles vaccination. Cyangugu: 11 indicators: same 5 indicators plus: family planning re-attendants, IUD implant and permanent FP methods as well as antenatal care and bed net distribution
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Many indicators are reported; it is not easy to assess their completeness and how denominators have changed over the study period
Selective reporting (reporting bias)	Unclear risk	With so many indicators, it is hard to rule out selective reporting
Other bias	High risk	Researchers involved in designing intervention and might therefore be biased
Baseline measurement	High risk	Not done
Matched characteristics for control study sites	Unclear risk	Not adequately described
Protection against contamination (intervention and controls)	Unclear risk	Not adequately described
Reliable primary outcome measure	Unclear risk	This cannot be ruled out, given the nature of the data collection (routine sources, not independently assessed)
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Soeters 2008

Methods	CBA - Democratic Republic of Congo Two districts in South Kivu were funded by a NGO to introduce PBF in 2006. Another 2 received input-based funding (drugs, supplies and fixed bonus payments for staff) from another NGO
Participants	Facilities - public, not-for-profit and private

Soeters 2008 *(Continued)*

Interventions	Conditional cash transfers with variable top-up for quality
Outcomes	16 indicators for health centres; 22 for hospitals; 8 for HIV/AIDS. Covered preventive care (ANC, FP, immunisation, bed nets, latrines, vitamin A) as well as tests (for TB), procedures (deliveries, IPD, OPD, treatment of STDs), outcomes (TB cases treated, sputum negative) and referrals to hospital. Also 3-monthly review of professional practice (checklist with 154 quality indicators) - 15% bonus if all met, or proportionately less as score declined
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not a cohort study so attrition is not relevant
Selective reporting (reporting bias)	Unclear risk	With so many indicators, it is hard to rule out selective reporting
Other bias	High risk	Researchers involved in designing intervention and might therefore be biased
Baseline measurement	Low risk	Done
Matched characteristics for control study sites	Unclear risk	Not adequately described
Protection against contamination (intervention and controls)	Unclear risk	Not adequately described
Reliable primary outcome measure	Low risk	Outcome measures reported in household studies
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Soeters 2009

Methods	CBA - Burundi PBF approach in 2 provinces; input-based funding in 2 controls. "Fund holder organizations" were established in the 2 intervention districts, which negotiated contracts with individual health facilities and verified and paid for the performance of the facility
Participants	Primary care facilities in 2 districts
Interventions	Conditional cash transfers, with added variable bonus for quality
Outcomes	23 output indicators established, covering preventive care; management of conditions; patient education. 16 used for hospitals

Soeters 2009 *(Continued)*

Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not done
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No information given on drop-out rates for household survey
Selective reporting (reporting bias)	High risk	With so many indicators, it is hard to rule out selective reporting. Not clear how the choice of indicators to include in the household survey was made (not all performance areas are covered).
Other bias	High risk	Researchers involved in designing intervention and might therefore be biased
Baseline measurement	Unclear risk	Done
Matched characteristics for control study sites	High risk	Considerable differences noted between intervention and control areas. No description given of how facilities were selected.
Protection against contamination (intervention and controls)	Unclear risk	Not adequately described
Reliable primary outcome measure	Low risk	Outcome measures reported in household surveys
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Vergeer 2008

Methods	CBA - Zambia Intervention facilities paid for meeting targets (50% of funding is fixed; 50% conditional on targets). 4 targets used, each with equal weight and set at same level for all facilities
Participants	Mission hospitals and health centres
Interventions	Target payments
Outcomes	4 indicators: inpatient turnover rate per bed (target 50); hospital deliveries per 1000 population (target 15); VCT user rate per 1000 population (target 15); drug availability measured by stock-out days
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement

Vergeer 2008 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not done
Incomplete outcome data (attrition bias) All outcomes	High risk	Many indicators are reported; it is not easy to assess their completeness and how denominators have changed over the study period
Selective reporting (reporting bias)	Unclear risk	Due to the large number of outcome measures, we cannot exclude the possibility of selective reporting
Other bias	Unclear risk	None identified
Baseline measurement	High risk	Not done
Matched characteristics for control study sites	Unclear risk	Sites chosen retrospectively and non-randomly; evidence of considerable differences
Protection against contamination (intervention and controls)	Unclear risk	No information provided
Reliable primary outcome measure	High risk	Outcome measure not independently assessed
Overall assessment of risk of bias	High risk	Our overall assessment was that there was a high risk of biased results from this study

Note on abbreviations: ANC = antenatal care; CBA = controlled before-after; FP = family planning; HC = health centre; HIV/AIDS = human immunodeficiency virus/acquired immune deficiency syndrome; IPD = inpatient day; IUD = inter-uterine device; OPD = outpatient visit; NGO = non-governmental organisation; PBF = performance-based financing; STD = sexually transmitted diseases; TB = tuberculosis; VCT = voluntary counselling and testing.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Bhushan 2002	This is a primarily a study of contracting in and out of health services, rather than paying for performance. The boundaries between the 2 are not always clear, but as a recent systematic review of contracting out exists, studies focusing on contracting have not been included
Biai 2007	Payment was not contingent on specific target, but a monthly sum for working in a ward implementing guidelines, filling in forms and being monitored
Chee 2004	This study evaluated the impact of GAVI-funding on vaccine coverage. The first 2 years of funding from GAVI were not based on performance. Only subsequent disbursements are based on performance (i.e. achieving increased immunisation coverage). The study does not attempt to evaluate the effect of only the performance-based component. Also, it seems impossible to disentangle the impact of the large funding that is awarded for the first 2 years, from the performance-based financing that is awarded later.
Chee 2007	This study evaluated the impact of GAVI-funding on vaccine coverage. The first 2 years of funding from GAVI were not based on performance. Only subsequent disbursements are based on performance (i.e. achieving increased immunisation coverage). The study does not attempt to evaluate the effect of only the performance-based component. Also, it seems impossible to disentangle the

Study	Reason for exclusion
	impact of the large funding that is awarded for the first 2 years, from the performance-based financing that is awarded later.
Chee 2008	This study evaluated the impact of GAVI-funding on vaccine coverage. The first 2 years of funding from GAVI were not based on performance. Only subsequent disbursements are based on performance (i.e. achieving increased immunisation coverage). The study does not attempt to evaluate the effect of only the performance-based component. Also, it seems impossible to disentangle the impact of the large funding that is awarded for the first 2 years, from the performance-based financing that is awarded later.
Eichler 2007	Although there are before-and after data for NGOs that did and did not participate in the PBF scheme, the groups cannot be assumed to be similar, particularly since criteria for inviting NGOs to participate were “indicators of readiness”
Furth 2006	Descriptive study of PBF, with interesting conclusions, but no control and so ineligible on study design grounds
Gauri 2004	No baseline data
Keller 2001	Evaluation study of performance based contracting project in Cambodia
Lim 2008	The study design is not clear although all other criteria have been met
Liu 2005	ITS, but only 2 time points before and after
Lu 2006	This study evaluated the impact of GAVI-funding on vaccine coverage. The first 2 years of funding from GAVI were not based on performance. Only subsequent disbursements are based on performance (i.e. achieving increased immunisation coverage). The study does not attempt to evaluate the effect of only the performance-based component. Also, it seems impossible to disentangle the impact of the large funding that is awarded for the first 2 years, from the performance-based financing that is awarded later.
Mahmood 2007	Simple before and after study
Marek 1999	Time-series data presented in Figures 4 and 5, but data collection started as the intervention programme was launched
Meessen 2006	Reviewed the institutional set-up of the PBF intervention, not the actual performance of the intervention
Meuwissen 2006	Looks at the changes in the doctors’ knowledge, attitudes and practice, which could be a result of the capacity-building sessions. Does not necessary look at the impact of the financial incentive. A before and after study, without control group
Schwartz 2004	This is a report on outcomes that is part of a larger body of evidence regarding the effects of contracting of health services in Cambodia
Soeters 2003	Simple before-after measurements
Sondorp 2009	There are measurements from 2 different time periods for areas that are being compared, but unclear whether the first is 'before' or early in the intervention period
World Bank 2003	Time series presented with 3 measures before and after introduction of the scheme in Burkina Faso. Not clear that the mechanism is PBF, as no rewards or sanctions are described
Yip 2001	CBA for change in payment system from fee for service to prepayment mechanisms for hospitals in Hainan. Changing provider payments studied, rather than PBF as such

CBA = controlled before-after study; GAVI = Global Alliance for Vaccines and Immunisation; ITS = interrupted time series; NGO = non=governmental organisation; PBF = performance-based financing

Characteristics of studies awaiting assessment *[ordered by study ID]*

Huntingdon 2010

Methods	Controlled before-after study
Participants	Primary healthcare units in Egypt
Interventions	Incentive payments compared to salary top-ups
Outcomes	Quality improvements in family planning, antenatal care and child health services
Notes	

Lundberg 2007

Methods	Cluster randomised trial
Participants	Private not-for-profit primary health care facilities in Uganda
Interventions	Three arms: control; increased financial autonomy; increased financial autonomy with target payments
Outcomes	Number of outpatient consultations; children immunised; antenatal care visits; attended births; malaria treatment for children; modern family planning uptake. Also perceptions of quality by patients and changing wealth profile of service users.
Notes	

ADDITIONAL TABLES

Table 1. Participants, scale of intervention, context and payment mechanism

Main studies	Participants	Scale of intervention	Context	Clinical group targeted	How payments were made	How were targets measured?
Peabody 2010	Public district hospitals	30 district hospitals; 617 doctors; 6000 children followed up	Philippines Facilities chosen from 30 districts in 11 provinces of the Visayas, a group of islands, and Mindanao Rural setting	Children under 6; inpatient care	Direct bonuses to doctors, paid quarterly, if quality scores for facility exceed cut-off (group 1); group 2 only motivated indirectly via Phil Health higher reimbursements to their facility	Quality of care was based on a combination of vignette scores (focusing on diarrhoea and pneumonia), case load and patient satisfaction

Table 1. Participants, scale of intervention, context and payment mechanism *(Continued)*

Basinga 2010	Primary care facilities	166 facilities in 16 districts (out of 401 primary care facilities in country)	Rwanda Outpatient care in mixed settings	Pregnant women and children under 6	Monthly payment per indicator, which is multiplied by quality index (0 to 1). Quality measures include inputs (staffing, drugs etc.) and processes	Facilities report to district committee each month; verified by periodic spot checks and auditing. Quality measured through quarterly routine visits by district supervisors.
Soeters 2009	Primary care facilities in 2 districts	40 primary facilities (30% faith-based; rest public). 10 subcontracts signed with private pharmacies. 2 intervention districts (Bubanza and Cankuzo) and 2 controls (Karuzi and Makamba)	Burundi Attempt to match intervention and control provinces. 2 said to be very poor; other 2 less so, though no details are given. Both supported by international NGOs, though in different ways.	General population	Fixed amount paid per targeted action; top-up of 15% available, based on quarterly reviews of quality (using composite index with 153 indicators)	Not clearly stated - presumably self reported by facilities, with some verification by fund holders
Canavan 2008	Mission hospitals, health centres and dispensaries in 5 dioceses	Total of 13 hospitals, 12 health centres and 29 dispensaries in PBF project; estimated population of 2.25 million	Tanzania Limited description of setting - assumed to be mix of rural and urban	General population	50% of support paid upfront for the year; 50% paid retrospectively if all the targets are hit (reviewed every 6 months)	Self assessed by facilities
Vergeer 2008	Mission hospitals and health centres	64 mission facilities supported in general, with catchment of 2 million people. However PBF element introduced to 5 hospitals and, later 3 health centres	Zambia Limited description of setting - assumed to be mix of rural and urban	General population	50% of support paid upfront for the year (fixed); 50% paid retrospectively if all the targets are hit	Self assessed by facilities
Soeters 2005	Public health centres only in Butare, but in Cyangugu also district hospitals	In Butare, 19 health centres, and target population of 380,000. In Cyangugu, 4 district hospitals, 25 health centres, 21 dispensaries (sub-contracted) and 15 health posts (also sub-contracted) - overall target population of 624,000	Rwanda No area characteristics given	General population, but with a strong focus on reproduc-	District health management teams paid to facilities - no description of frequency	Reported by facilities but periodic verification by patient surveys (by community groups in Cyangugu, public health school in Butare)

Table 1. Participants, scale of intervention, context and payment mechanism *(Continued)*

				tive health		
Soeters 2008	Facilities - public, not-for-profit and private	Two intervention districts had populations of 300,000 and neighbouring controls had 232,000. In 2 intervention districts, fund holder signed with 39 health centres and community groups (for monitoring) and 4 hospitals; health centres in turn could sub-contract to private clinics (22 did so)	Democratic Republic of Congo Mixed setting	General population	Fixed amount paid per targeted action per month; top-up of 15% available, based on quarterly reviews of quality (using composite index with 154 indicators). Also 15% additional payment for remote facilities. Population coverage targets established but do not appear to have been used in relation to payments	Reported by facilities but periodic verification by contracted community monitoring groups
Liu 2003	Doctors	All 127 county general hospitals in Shandong were surveyed (covering a population of 86 million). 108 responded.	China Hospital care in mixed rural/urban area (75% rural inhabitants)	Unspecified	Unspecified. Assumed to be added monthly to pay packets	Not specified
Quy 2003	Private doctors and pharmacists (latter only able to refer, not diagnose and treat)	185 doctors and 150 pharmacists in 2 (out of 22) districts of Ho Chi Minh City	Vietnam Ambulatory care in urban and semi-urban area	TB smear-positive patients	Not stated	Based on routine reporting forms provided by the National TB Programme

NGO: non-governmental organisation; PBF: performance-based financing; TB: tuberculosis

Table 2. Magnitude of incentives and ancillary components

Main studies	Who received payments?	Magnitude of incentives paid	Relative size of incentives	Additional to ordinary wages/ funding?	Source of funding	Any ancillary components?	Overall cost	Comments on targets and target setting
Peabody 2010	Bonuses shared 50% between physicians and other staff (nurses, administrators, support staff)	PhP 100 per patient per day	5% (2% to 8%) of doctor's income	Yes	Phil Health (public body) paid for intervention; US National Institutes of Health paid for research	All groups got quality assessment and feedback. Only intervention groups received additional funding.	Not given	Based on clinical protocols
Basinga 2010	Facilities receive funding and have discretion over its use	Some examples: USD 4.59 paid per delivery. USD 0.18 paid per child preventive visit. USD 1.83 for referral of malnourished child. USD 0.92 for TT and malaria prophylaxis during ANC. USD 0.09 per ANC visit (USD 0.37 for all 4 visits).	Facility funding increased by 22% for PBF group as a whole. On average facilities allocated 77% of the PBF funds to increase personnel compensation, amounting to a 38% increase in staff salaries	Yes	Government of Rwanda pays, but funded from pooled fund, with donor support (also external support for the evaluation)	Additional funding, but controlled for by paying equivalent to input-based group. Supervisors from the District Hospital would, at the end of their visits, discuss their findings with the facility's personnel, and provide recommendations to improve quality of services (part of routine supervision for all facilities). Increased monitoring and evaluation. Payments also fed through into increased staff payments	Not stated Cost-effectiveness estimates due out later	Indicators set by MoH, based on priorities, available budget and previous NGO experience
Soeters 2009	Facilities and, via facility managers, staff. Bonuses paid to staff based on performance criteria such as responsibility of staff, years of employment,	Ranged from USD 0.25 per new OPD case to USD 10 for TB diagnosis	Not clear in relation to salaries, but for facilities, in 2008, PBF payments constituted 58% of facility total revenue (PBF,	Yes	International NGO, Cordaid, funded by Dutch government and EU	Increased funding; management support to facilities, helping them to develop business plans and to manage finances, including how to calculate staff bonuses. In addition, the fund holder could apply a specif-	The package of subsidy was USD 2 per person, with USD 0.6 added for management etc. On top of	Based on early discussions of stakeholders

Table 2. Magnitude of incentives and ancillary components *(Continued)*

	overtime or hours worked as well as individual performance reviews		government payments, user fees) in surveyed facilities in intervention areas			ic isolation bonus for each health facility of between 5% and 40% based on criteria such as isolated location, an extremely poor population, the state of the infrastructure or when there are many displaced persons in the catchment area. Workshops were held to train and involve a range of stakeholders, from 2006-9	that USD 1.70 came from the Ministry of Health to fund free care	
Canavan 2008	Cordaid sets guidelines for use of bonuses: 50% maximum for staff motivation; equipment, drugs and supplies to a maximum of 30%; infrastructure max 20%; running cost (including maintenance and communication) to a maximum of 10%. The District Health Offices are eligible for 25% of the performance bonus allocation	Overall ceilings were 42 million Tsh per hospital; 14 million Tsh per HC; 7 million Tsh per dispensary	8% of facility income on average (for all Cordaid support)	Yes	International NGO: Cordaid	In addition to funding, Cordaid has supported the diocese for decades with human and financial support	Budget set at 0.5 Euros per capita	Indicators agreed with diocese, not facilities, and do not reflect baseline states or operational constraints. Also too focused on curative care. Authors also note absence of routine data for some of the facilities, and the fact that population denominators vary (and in some cases are based on norms rather than actual figures), thus rendering indicators unreliable
Vergeer 2008	Payments made to facilities but Cordaid specified that	Maximum of 90 million ZMK per hospital and 50 million	17% of facility revenue. Amounts small rela-	Yes	International NGO Cordaid	15,000 Euros given to HCs and 30,000 Euros to hospitals to allow them to rehabilitate or	Budget set at 0.5 Euros per capita	Lack of involvement of other stakeholders - authors recommend integrating planning within

Table 2. Magnitude of incentives and ancillary components *(Continued)*

	40% to 60% of payments could go to staff. Rest for infrastructure, drugs, supplies and running costs (with guidelines for allocation to different categories)	ZMK for health centres paid for meeting targets. Maximum of 20% paid on top for District Health Office management costs	tive to salaries (but variable by facility).			buy supplies at start of project. Authors note need for more technical support and capacity building in implementing project.		district. The current indicators used, all point to quantitative aspects of health service delivery and do not include performance measures on the quality of health services which Cordaid also aims to improve. Nor do the indicators selected reflect preventative aspects of health care provided, crucial for the provision of integrated health care. The payment for performance depends on a uniform target set for each indicator rather than based on baseline or contextual circumstances like population catchment or available staffing. As a result, one health facility would need to perform much harder than others to receive the same amount of incentives. Contracts were signed with the diocese, not the facilities. Facilities were not involved in discussion of indicators, and were sometimes not even aware of them
Soeters 2005	Facilities. Some funds passed on to staff in the form of bonus, but the mechanism for this is not described	Payment per specified services ranged from F40 for new OPD to F2,500 for facility delivery (in Butare); in Cyangugu from F150 for new	50% and 70% of external support to the facilities (funds not raised directly from the community) came from the PBF in Butare and	Yes	Healthnet International and Cordaid	Increased funding; technical support	Total spending in one province was USD 1.5 per capita; other province figure not given but said to be lower	Not described

Table 2. Magnitude of incentives and ancillary components *(Continued)*

		OPD to F15,000 for CS	Cyangugu respectively					
Soeters 2008	Facilities, but discretion to pass on to staff	Payments for health centres ranged from USD 0.30 per bed-day to USD 20 per TB patient successfully treated. Total of USD 200 to 4,000 per facility per month	Not stated, but the incentives must have been the major component of funding for the health centres at least, as their overall revenues per person rose from USD 0.51 to 1.04 over the period, while the budget for the PBF was USD 2 per capita. No state support is received in most cases	Yes	Cordaid supported intervention; IRC the controls. IRC funded by DFID. Cordaid funding not mentioned	Each facility supported to develop a business plan every 3 months. In addition, a fixed amount of USD 2,000 per month was paid to the district team to supervise (so that they would not tax the facilities)	USD 2.4 paid per capita in intervention zones (USD 2 for basic package and USD 0.4 for administrative expenses and overheads); USD 9 to 12 per capita was spent in control districts	Not clear how the amounts per indicator were fixed
Liu 2003	Doctors	Not stated	Not stated	Yes	Not stated but came from general facility revenues	Not applicable	Not specified	Set by hospital administrators
Quy 2003	Doctors	USD 0.80 paid for each sputum smear-positive case detected, USD 10 for each case treated (whatever the outcome), and USD 5 per case transferred to	Not stated	Yes	Government. Study funded by SIDA/SARED/AHPSR	National guidelines provided, possibly with some training. Referral forms also provided, and quarterly supervision visits made	Not stated	The targets are routine TB indicators

Table 2. Magnitude of incentives and ancillary components *(Continued)*

	the National TB programme

Note on abbreviations: ANC = antenatal care; CS = Centre de Sante (Health centre); DFID = Department for International Development (UK aid); F = Rwanda Franc; HC = health centre; IRC = International Rescue Committee; MoH = Ministry of Health; NGO = non-governmental organisation; OPD = outpatients department; PBF = performance-based financing; PhP = Philippines Peso; SIDA = Swedish International Development Agency; TB = tuberculosis; Tsh = Tanzanian Shilling; TT = tetanus toxoid; ZMK = Zambian Kwacha

Table 3. Reporting of outcomes in included studies

Studies	Health professional outcomes/process measures	Patient outcomes	Economic variables (e.g. change in resource use)	Length of time during which outcomes were measured after initiation of intervention	Measurement of outcome indicator
Peabody 2010	Quality of care, as scored by vignettes, covering 5 domains (history taking, physical examination, test ordering, diagnosis and treatment) administered to random sample of doctors in each hospital every 6 months. Accompanied by physician survey	4 health status measures were evaluated at baseline (2003) and 2 years post-intervention (2007) in children under 5 hospitalised for diarrhoea and pneumonia: (1) age-adjusted wasting, (2) C-reactive protein in blood, (3) haemoglobin and (4) general self reported health (GSRH), the parental assessment of the child's health	Cost of care to patients, assessed by exit interviews	6 to 36 months	Health outcomes assessed by independent party 4 to 10 weeks post-discharge
Basinga 2010	Antenatal visits; 4 or more visits per pregnant woman; institutional deliveries; visits by children (under-2s; and 2 to 5); quality of ANC care; TT vaccine administered during ANC visit; fully immunised children < 2 years	Further paper due out on this in due course	Not reported	18 to 24 months	Assessed through baseline and second round survey of facilities and households. Also exit interviews
Soeters 2009	Changes in utilisation (institutional deliveries; ANC; treatment episodes; immunisation)	Household behaviour and awareness variables (having bed net; using FP; treating child with diarrhoea with ORS; heard about condoms; heard about HIV). No outcome variables	Household out of pocket payments; health centre revenues; staff pay	2 years (2006-8)	Household survey carried out pre- and post (500 households). Quality reviews done of 29 facilities, and interviews held with managers of same facilities.
Canavan 2008	Change in coverage indicators: VCT per population; IPD per population; OPD per population; institutional delivery per population; % availability 10 essential drugs	Not reported	Not reported	1 year	Self reported by facilities
Vergeer 2008	Change in coverage indicators: VCT per population; IPD per population; OPD per population; institutional delivery per population; % availability 10 essential drugs	Not reported	Not reported	1 year	Self reported by facilities

Table 3. Reporting of outcomes in included studies *(Continued)*

Soeters 2005	Utilisation measures - outpatient visits, FP uptake, ANC, immunisation, supervised deliveries	Not reported	Change in revenue for health centres	2 years	Self reported by facilities
Soeters 2008	7 output and 2 patient knowledge indicators tracked	None reported	Changes to revenues at the facilities and payments by households	2 years	Household survey (pre/post); facility survey; interviews with managers
Liu 2003	Not reported	Not reported	Change in hospital revenue is the only outcome measure	4 years before and 3 years after switch	Questionnaire sent to hospital administrators
Quy 2003	Case detection was the main outcome indicator	Not reported	Not reported	9 months	Routine NTP forms

Note on abbreviations: ANC = antenatal care; FP = family planning; OPD = outpatients department; ORS = oral rehydration solution; NTP = National Tuberculosis Programme; TT = tetanus toxoid; VCT = voluntary counselling and testing

Table 4. Summary of findings for primary outcomes

Studies	Changes in the utilisation of health care services	Quality of care	Changes in patient outcomes	Unintended effects	Changes in resource use
Peabody 2010	"By measuring the average number of monthly inpatients, we found that there was no substantive change in patient volumes in the hospitals that qualified for bonuses (329) versus those that did not (347)". Volumes did increase, as expected, in the insurance group	All hospitals started at 66% quality score. PBF group scored 76% on vignettes (after 36 months), compared with 72% for control. Gains statistically indistinguishable from the improvement in the insurance group. Disaggregating the quality score, there were gains across all 5 domains of quality. Con-	Among children, many of whom had poor baseline health status, compared to controls, the intervention sites showed improvements of 7% and 9% points in GSRH and wasting, respec-	Not reported	Investigated relationship between quality scores and cost of care in health insurance and bonus (PBF) group. Found that increase from low to average score was associated with 45% drop in cost of care for patients - after that costs started to rise

Table 4. Summary of findings for primary outcomes *(Continued)*

		trol group also improved (dissemination effect) but not as much as the 2 intervention arms (both about 10% improvement over the period)	tively (P ≤ 0.001). Health status in control group got worse over the period, presumably due to exogenous factors. No analysis for health insurance arm. No significant difference for anaemia and CRP		
Basinga 2010	Any prenatal care: absolute risk increase of 0.002 (- 0.022 to 0.025). Relative risk increase (all RR judged relative to baseline levels for intervention group*): 0.002. Made 4 or more prenatal visits: increased absolute risk 0.01 (-0.063 to 0.083); relative risk increase 0.056. Institutional deliveries: absolute risk increase 0.074 (0.006 to 0.142); relative risk 0.211. Child preventive visit (0 to 23 months):absolute risk increase 0.134 (0.045 to 0.224); relative risk increase 0.638. Child visits (24 to 59 months): absolute risk increase 0.106 (0.05 to 0.161); relative risk increase 1.325. Child (12 to 23 months) is fully immunised (based on cards): absolute risk increase -0.065 (-0.178 to 0.047); relative risk increase -0.105.	Quality of prenatal visit higher for intervention group: 0.14 (0.015 to 0.265). Tetanus vaccine during prenatal visit: absolute risk increase of 0.054 (0.007 to 0.10); relative risk increase 0.076.	Not reported	Not reported	PBF payments increased average overall facility expenditure by 22%
Soeters 2009	Found significantly better performance for institutional deliveries in intervention group and proportion of households having at least 1 bed net, but no significant difference for other output indicators. The proportion of episodes treated in modern health facilities in the PBF provinces increased in 2008 to 95% and to 92% in the control provinces, which indicates that there were few	There was an improvement in the PBF provinces for quality of care, as assessed by health professionals in hospitals and health centres	There were no significant differences in the 4 household knowledge indicators	Not reported	Average household out of pocket payment increased in control areas compared to intervention, though was much lower before (and still lower in absolute terms afterwards). Proportionate to income, it increased in control areas, especially for poorest group. Number not seeking care for financial reasons fell in both groups. Payment per primary care vis-

Table 4. Summary of findings for primary outcomes *(Continued)*

	barriers for patients to attend modern health facilities.	(36% and 50% improvement respectively), while it slightly reduced in the control provinces by 13% (hospitals) and 18% (health centres)			it rose in controls and fell in intervention, though very different starting points (higher in intervention). Cost for delivery fell in both groups, though was much higher to start with in intervention group and fell furthest there. In health centres, overall revenues (combining government funding, donor funding and user fees) rose in intervention group from USD 0.6 to 1.55 per capita; in intervention group fell from USD 0.65 to 0.57 per capita. This mainly reflects changes to donor funding and government (both increased in intervention and reduced in controls). Staff pay also rose significantly in intervention areas (158% increase for heads of facility, for example), while declining in intervention provinces. Number of staff also rose in intervention areas than in controls.
Canavan 2008	1. Decline in IPD in mission facilities (greater than small decline in government facilities). 2. Decline in institutional deliveries at mission health centre, relative to increase for government facilities. In hospitals, both declining. 3. Increase in VCT, but much smaller than government facilities, where this doubled (but based on sample of 2 facilities). 4. ANC: greater decline for mission facilities at dispensary level, compared to government facilities, and smaller (slight) increase at HC level. Both very high levels (unlike in hospitals, where lower and declining for both groups). 5. For OPD at HC level, utilisation below national target, and similar trend for both groups (increase of around 0.1 in new contacts per pop.). For dispensaries, increase in mission facilities, but absolute level much lower level than government facilities	Not reported	None reported	Authors comment that it may cause focus on curative care, as there were no preventive indicators or indicators linked to quality of care	Drug supplies improving over 2004-7 for both groups. Authors comment that there were some delays in paying bonus funds. Managers also asked for more flexibility in use of funds
Vergeer 2008	1. Increase in VCT across both groups, but this is also affected by other donor investment at this time, 2. Institutional deliveries stable overall for both groups, at hospital and HC level, 3. No significant change in ANC, 4. No significant difference in intervention and control hospitals in relation to IPD/OPD. Variety of patterns across facilities	Not reported	None reported	Increasing the inpatient turnover rate especially appeared questionable for health	Improvements in drug supplies over 2004-7, for intervention and control groups. Delays in payments reported. Size of transaction costs not clear

Table 4. Summary of findings for primary outcomes *(Continued)*

				centres which are to focus mainly on preventative and promotive health care; managers felt it led to a neglect of PHC. TB detection rates were investigated to see if they fell in PBF areas - there was no clear pattern	
Soeters 2005	OPD: increase in OPD by 147% in intervention zone, compared with 52% in control. Highest in province with higher tariffs. Still low in absolute terms - 0.55 per person per year. For institutional deliveries, increase of 10.9% versus 2.9% for controls. Still low in absolute terms - 23%. For FP, intervention achieved 2.8% increase in new acceptors, relative to 0.2% increase for control, though again very low absolute coverage (3.9%). For measles immunisation, coverage increased by 11% in intervention, versus 1% in controls, though controls were higher to begin with (therefore harder to increase); overall high coverage of over 80%	Intervention facilities higher score for checklist of quality of care (75%, versus 47% for controls), but no baseline to compare with	Not reported	Not reported	Lower total revenues in intervention zones before, but higher after (400% increase in external funding over period for intervention provinces; 150% increase for controls). Community contributions decreased in both zones, but more so in control (-7%, compared with -2% in intervention zone). 72% of health workers in intervention zones report pay increases as a result of the PBF. Pay of trained health workers 30% higher in intervention provinces (no baseline data but this is attributed to the bonuses - 95% of PBF revenues go to staff in Butare; 40% in Cyangugu). Untrained staff numbers found to be 35% higher in intervention provinces
Soeters 2008	Only two indicators had significant difference ($P < 0.05$) between control and intervention group: "ever heard of HIV" was 10% more likely in PBF, and institutional delivery was 29% more likely in control	Not reported	Not reported	Not reported	Household payments rose by 45% over period in PBF areas, compared to 18% drop in controls. 6.7% of total HH expenditure went on health care in 2008 in PBF areas, compared to 5.45% in controls (drop of 3.5% in PBF areas

Table 4. Summary of findings for primary outcomes *(Continued)*

					over the period and 4.5% in controls). Annual per capita user fee revenue increased in PBF health centres by 25% over 2005-8, while it fell by 43% in controls, where user fees reduced (in PBF collected USD 0.64, compared to USD 0.19 in controls). Annual per capita external assistance rose from USD 0 in 2005 to USD 0.36 in PBF health centres, while in controls it went from USD 0.17 to 0.31; in total, revenues of USD 1.04, compared to USD 0.45 in controls). Significant increase in facilities meeting staffing norms in intervention facilities, compared to controls (P < 0.04)
Liu 2003	Not reported	Not reported	Not reported	Not reported	ITS reanalysis found a significant change in level and slope for group A (switch from no bonus - salary only - to flat bonus). All other comparisons not significant (see reanalysis table). This differs from the authors' conclusions that the switch to high powered incentives led to "a consistent sudden increase in the rate of growth of hospital revenues". Patterns of revenue growth varied considerably across years, with often an increase after the switch, followed by a fall back in growth of revenues. The cross-sectional regression analysis for 1997 supported the importance of the bonus type, with revenue-related bonuses associated with significantly higher overall revenues than quantity-related ones.
Quy 2003	Not reported	No significant shift in level or slope, according to ITS reanalysis (see reanalysis table). This differs from the authors' findings that "The case detection of new sputum smear-positive cases in PPM districts increased by 18% (21/100	Not reported	Not reported	Not reported

Table 4. Summary of findings for primary outcomes *(Continued)*

		000, 95%CI 0–42) compared to the previous year, while a slight decrease occurred in control districts"		

Note on abbreviations: ANC = antenatal care; CRP = c-reactive protein; FP = family planning; GSRH = general self reported health; HC = health centre; IPD = inpatient day; HH = household; ITS = interrupted time series; OPD = outpatients department; PBF = performance-based financing; PHC = primary health care; PPM = public private mix; TB = tuberculosis; VCT = voluntary counselling and testing
*Note that the relative risk increase is calculated by applying the estimated absolute effect to the baseline level in the intervention group.

Table 5. Utilisation changes, hospitals, Canavan 2008

IPD Admissions	**2005**	**2006**	**2007**	Change 05-07	***Direction of change***
Rubya Hospital (M)	18501	17632	15517	-16.13%	*Down*
Isingiro Hospital (M)	1034	670	687	-33.56%	*Down*
Average intervention	9768	9151	8102	-17.05%	
Nyakahange Hospital (G)	15902	15713	15899	-0.02%	*Unchanged*
ELCT Nodolage Hospital	6747	7148	7221	7.03%	*Up*
Average control	11325	11431	11560	2.08%	
Number of pregnant women having 4 or more ANC visits in health facility	**2005**	**2006**	**2007**	Change 05-07	
Rubya Hospital (M)	2078	1641	1691	-18.62%	*Down*
Isingiro Hospital (M)	162	254	240	48.15%	*Up*
Average intervention	1120	948	966	-13.79%	
Nyakahange Hospital (G)	1599	1639	1687	5.50%	*Up*
ELCT Nodolage Hospital	792	635	733	-7.45%	*Down*
Average control	1196	1137	1210	1.21%	
Number of pregnant women with TT3	**2005**	**2006**	**2007**	Change 05-07	
Rubya Hospital (M)	1793	1272	1399	-21.97%	*Down*
Isingiro Hospital (M)	683	249	386	-43.48%	*Down*
Average intervention	1238	761	893	-27.91%	

Table 5. Utilisation changes, hospitals, Canavan 2008 *(Continued)*

Nyakahange Hospital (G)	1082	822	841	-22.27%	*Down*
ELCT Nodolage Hospital	590	469	572	-3.05%	*Down*
Average control	836	646	707	-15.49%	
Number of deliveries that were professionally attended - institutional deliveries (hospitals)	**2005**	**2006**	**2007**	Change 05-07	
Rubya Hospital (M)	2974	2737	2712	-8.81%	*Down*
Isingiro Hospital (M)	683	731	615	-9.96%	*Down*
Average intervention	1829	1734	1664	-9.02%	
Nyakahange Hospital (G)	1927	1755	1768	-8.25%*	*Down**
ELCT Nodolage Hospital	1927	1755	1768	-8.25%*	*Down**
Average control	1927	1755	1768	-8.25%*	
No. of children with DPT3	**2005**	**2006**	**2007**	Change 05-07	
Rubya Hospital (M)	1862	1633	990	-46.83%	*Down*
Isingiro Hospital (M)	668	399	595	-10.93%	*Down*
Average intervention	1265	1016	793	-37.35%	
Nyakahange Hospital (G)	1495	1728	1231	-17.66%	*Down*
ELCT Nodolage Hospital	570	572	518	-9.12%	*Down*
Average control	1033	1150	875	-15.30%	
Total no. of new consultations	**2005**	**2006**	**2007**	Change 05-07	
Rubya Hospital (M)	17366	17117	15307	-11.86%	*Down*
Isingiro Hospital (M)	4058	6145	6022	48.40%	*Up*
Average intervention	10712	11631	10665	-0.44%	
Nyakahange Hospital (G)	51868	35477	35910	-30.77%	*Down*
ELCT Nodolage Hospital	8700	8158	9110	4.71%	*Up*
Average control	30284	21818	22510	-25.67%	

*Likely to be a data error, as numbers are reported as being the same for both 'control' hospitals

Table 6. Utilisation changes, health centres and dispensaries, Canavan 2008

Number of pregnant women having 4 or more ANC visits in HF	**2005**	**2006**	**2007**	**Change 05-07**

Table 6. Utilisation changes, health centres and dispensaries, Canavan 2008 *(Continued)*

Mwemage HC (M)	331	300	339	2.42%	*Up*
Rwanbiazi HC (M)	876	730	1010	15.30%	*Up*
Average intervention	604	515	675	11.76%	
Nkwenda HC (G)	1958	1910	1753	-10.47%	*Down*
Kiagara HC (G)	3129	3243	3732	19.27%	*Up*
Average 'control'	2544	2577	2743	7.82%	
Number of deliveries that were professionally attended - institutional deliveries (health centres)	**2005**	**2006**	**2007**	**Change 05-07**	
Mwemage HC (M)	291	322	268	-7.90%	*Down*
Rwanbiazi HC (M)	113	331	202	79.00%	*see (1)*
Average intervention	202	326	235	16.41%	
Nkwenda HC (G)	461	551	615	33.41%	see (2)
Kiagara HC (G)	1020	1127	1785	75.00%	*Up*
Average 'control'	741	839	1200	62.05%	
Total no of new consultations (HCs)	**2005**	**2006**	**2007**	**Change 05-07**	
Mwemage HC (M)	2214	3957	3621	63.55%	*Up*
Rwanbiazi HC (M)	1804	2685	2655	47.17%	*Up*
Average intervention	2009	3321	3138	56.20%	
Nkwenda HC (G)	6342	5886	9880	55.79%	*
Kiagara HC (G)	6342	5886	9880	55.79%	*Up*
Average 'control'	6342	5886	9880	55.79%	
Total no of new consultations (dispensaries)	**2005**	**2006**	**2007**	**Change 05-07**	
Kishuro (M)	1399	1163	2101	50.18%	*Up*
Rwenkede (M)	290	769	882	204.14%	*Up*
Average intervention	845	966	1492	76.61%	
Kyerwa (G)	5124	1748	1660	-67.60%	*Down*
Buhembe (G)	10299	7441	7912	-23.18%	*Down*
Average control	7712	4595	4786	-37.94%	

Note on abbreviations: DPT = diphtheria, pertussis and tetanus; g = government facilities (non-intervention); m = mission facilities (intervention group)
*Probable data error as both hospitals have the same figures
(1) Data sheets contained decimals – not natural units
(2) Data were calculated, not recorded

Table 7. Utilisation changes, Vergeer 2008

Inpatients per consulting staff	**2006**	**2007**	*Change*	
Lubwe	812	566	-30.30%	Down
Kasaba	1474	1753	18.93%	Up
St. Paul's	1456	2234	53.43%	Up
Minga	435	846	94.48%	Up
Average PBF	1044	1350	29%	Up
Mbereshi*	976	916	-6.15%	Down
Petauke	321	335	4.36%	Up
Average non-PBF	648.5	625.5	-3.5%	Down
Institutional deliveries	**2006**	**2007**	*Change*	
Lubwe	440	469	6.59%	Up
Kasaba	214	234	9.35%	Up
St. Paul's	1825	1801	-1.32%	Down
Minga	533	978	83.49%	Up
Average PBF	753	871	15.6%	Up
Mbereshi	209	775	270.81%	Up*
Petauke	1172	1310	11.77%	Up
Average non-PBF	690.5	1042.5	50.98%	Up
Total no. outpatient consultations	**2006**	**2007**	*Change*	
Lubwe	10653	15823	48.53%	Up
Kasaba	12659	17830	40.85%	Up
St. Paul's	10147	14768	45.54%	Up
Minga	20011	21548	7.68%	Up
Average PBF	13368	17492	30.86%	Up
Mbereshi	11774	11235	-4.58%	Down

Table 7. Utilisation changes, Vergeer 2008 *(Continued)*

Petauke	14078	39675	181.82%	Up
Muzeyi	24388	25368	4.02%	Up
Chiparamba	14383	13823	-3.89%	Down
Katete	30060	36614	21.80%	Up
Average non-PBF	18937	25343	33.8%	Up
No of TB cases	**2006**	**2007**	*Change*	
Lubwe	157	189	20.38%	Up
Kasaba	142	86	-39.44%	Down
St. Paul's	247	351	42.11%	Up
Minga	364	459	26.10%	Up
Average PBF	228	271	19.23%	Up
Mbereshi	124	134	8.06%	Up
Petauke	213	324	52.11%	Up
Muzeyi RHC	69	80	15.94%	Up
Chiparamba RHC	21	24	14.29%	Up
Katete UHC	63	37	-41.27%	Down
Average non-PBF	98	120	22.24%	Up
Number of women under ANC (first attendance)	**2006**	**2007**	*Change*	
Lubwe	755	985	30.46%	Up
Kasaba	775	691	-10.84%	Down
Minga	1193	1282	7.46%	Up
Average PBF	908	986	8.63%	Up
Mbereshi	412	221	-46.36%	Down
Average non-PBF	412	221	-46.36%	Down
Number of pregnant women re-attending ANC	**2006**	**2007**	*Change*	
Lubwe	1277	1376	7.75%	Up
Kasaba	1354	1173	-13.37%	Down
Minga	2067	2371	14.71%	Up

Table 7. Utilisation changes, Vergeer 2008 *(Continued)*

Summary PBF	1566	1640	4.73%	Up
Mbereshi	598	405	-32.27%	Down
Average non-PBF	598	405	-32.27%	Down
Inpatient admissions	**Change 2006-2007**			
Lubwe	Little or none			
Kasaba	Little or none			
St. Paul's	Little or none			
Minga	Up			
Summary PBF	Mixed			
Mbereshi*	Little or none			
Petauke	Up			
Summary non-PBF	Mixed			
% deliveries ending in C-section	**Change 2006-2007**			
Lubwe	Up			
Kasaba	Little or none			
St. Paul's	Down			
Minga	Up			
Summary PBF	Mixed			
Mbereshi	Down			
Petauke	Up			
Summary non-PBF	Mixed			
% Post-operative wound infections	**Change 2006-2007**			
Kasaba	Little or none			
St. Paul's	Up			
Summary PBF	Mixed			
Petauke	Up			
Summary non-PBF	Up (N = 1)			
% basic drug stock available	**Change 2006-2007**			

Table 7. Utilisation changes, Vergeer 2008 *(Continued)*

Lubwe	Little or none
Kasaba	Little or none
St. Paul's	Little or none
Minga	Little or none
Summary PBF	Little or none
Mbereshi	Little or none
Petauke	Up
Summary non-PBF	Mixed

* suspected reporting error

Table 8. Utilisation outcomes with relative risks (five CBA studies)

				Baseline values	
Outcome	**Adjusted RR**	**Confidence interval**	**P value**	**Control**	**Intervention**
Soeters (Burundi)					
Institutional delivery	1.79	(1.38 to 2.32)	< 0.0001	73%	48%
Vaccinated according to schedule for child's age	1.12	(1.03 to 1.21)	0.01	88%	89%
Pregnant women completed vaccination	1.08	(0.74 to 1.57)	0.61	70%	70%
Coverage of modern family planning	2.04	(0.61 to 6.84)	0.25	8%	8%
Coverage of bed nets (1 person per net)	1.50	(1.16 to 1.93)	< 0.01	11%	17%
Coverage of bed nets (1.5 persons per net)	1.50	(1.17 to 1.94)	< 0.01	17%	25%
Four immunisation indicators under 1 fully satisfied	1.06	(0.95 to 1.18)	0.33	90%	91%
Child with diarrhoea treated with ORS	1.06	(1.03 to 1.08)	< 0.0001	87%	81%
Heard about condoms	0.94	(0.91 to 0.97)	< 0.0001	96%	96%
Heard about family planning	1.01	(1.00 to 1.02)	0.17	97%	98%
Household has latrine in good condition	1.14	(0.72 to 1.81)	0.56	30%	22%
Canavan (Tanzania)					
IPD admissions	0.82	(0.76 to 0.89)	< 0.0001	25%	14%

Table 8. Utilisation outcomes with relative risks (five CBA studies) *(Continued)*

Utilisations in health centres	0.94	(0.83 to 1.08)	0.40	20%	15%
Soeters (Rwanda)					
Family planning	1.26	(0.41 to 3.87)	0.68	0%	1%
Institutional delivery	1.08	(0.63 to 1.84)	0.79	7%	12%
Soeters (Democratic Republic Congo)					
Institutional deliveries	0.75	(0.67 to 0.84)	< 0.0001	60%	71%
At least 1 bed net in household	1.40	(0.68 to 2.87)	0.36	14%	12%
Vaccinated according to schedule for child's age	1.08	(0.83 to 1.41)	0.57	67%	58%
Pregnant women completed vaccination	1.13	(1.05 to 1.21)	0.00	65%	52%
Household with latrines in reasonable condition	0.96	(0.45 to 2.06)	0.92	8%	17%

Note: RRs are adjusted for baseline imbalances in outcome between intervention and control, plus adjusted for the clustering caused by introducing in some regions and not in others (RR > 1 favours the intervention, and RR < 1 favours the control)
*Measles vaccination in Rwanda was not included as the number of vaccinated was larger than the estimated target population for Bugurama in both 2001 and 2004.
CBA = controlled before-after study; IPD = inpatient day; ORS = oral rehydration solution; RR = relative risk

Table 9. ITS analysis of Quy 2003

	Immediate change in level		Slope		Change in slope	
Quy 2003	Estimate	P value	Pre intervention	Post intervention	Estimate	P value
S+ TB cases/100,000 person years - unadjusted for seasonality	5.4530	0.2227	-0.5791	-0.1773	0.4022	0.8936
S+ TB cases/100,000 person years - adjusted for seasonality	0.8717	0.7435	0.1063	3.3718	3.2655	0.1241

ITS = interrupted time series; TB = tuberculosis

Table 10. Summary of difference-in-difference estimates and relative risks of health outcomes comparing intervention and control, Peabody 2010

Health outcome	Pre-bonus intervention prevalence (%)	Post bonus intervention prevalence (%)	Absolute change (post-pre) %	Relative % change	P value

Table 10. Summary of difference-in-difference estimates and relative risks of health outcomes comparing intervention and control, Peabody 2010 *(Continued)*

CRP negative					
Intervention	97.69	98.07	0.38		
Control	96.06	95.60	-0.46		
Difference	1.63	2.47	0.84	0.86%	0.497
Not anaemic					
Intervention	93.80	91.95	-1.85		
Control	89.59	92.61	3.02		
Difference	4.21	-0.66	-4.87	-5.19%	0.253
Not wasted					
Intervention	70.09	69.57	-0.51		
Control	75.02	65.25	-9.77		
Difference	-4.93	4.32	9.25	13.20%	< 0.0001
GSRH at least good*					
Intervention	78.50	85.02	6.53		
Control	86.79	85.94	-0.85		
Difference	-8.29	-0.92	7.37	9.39%	0.001

Note: absolute figures from investigators; relative change added by review authors (absolute change/pre-bonus intervention prevalence in the intervention group) x 100
*GSRH (general self reported health) allowed parents to rate their child's health as excellent, very good, good, fair or poor: this was dichotomised to 'good or better' versus 'fair or poor'.
CRP = C-reactive protein

Table 11. ITS reanalysis of Liu 2003

Liu 2003		Immediate change in level		Slope		Change in slope	
Outcome	Comparison	Estimate	P value	Pre intervention	Post intervention	Estimate	P value
Average Service Revenue (yuan) in real terms (base year = 1975)	A	141,104	0.0176	56,840	163,586	106,747	0.0021
Average Service Revenue (yuan) in real terms (base year = 1975)	B	93,582	0.1680	132,116	157,499	25,383	0.2211
Average Service Revenue (yuan) in real terms (base year = 1975)	C	-108,376	0.7190	171,045	364,267	193,222	0.1650
Average Service Revenue (yuan) in real terms (base year = 1975)	D	218,375	0.3871	226,064	245,914	19,851	0.7842
Average Service Revenue (yuan) in real terms (base year = 1975)	E	-62,535	0.8812	160,735	308,496	147762	0.3687
Average Service Revenue (yuan) in real terms (base year = 1975)	F	399,048	0.6347	49,352	660,296	610,944	0.0672

Group A compared a "flat bonus" ("a certain percentage of the hospital's revenue was distributed almost evenly to all staff member every month if the revenue target of the hospital was reached") with a no bonus group. Group B compared a 'quantity-related bonus' with no bonus. Group C compared a 'revenue-related bonus' with no bonus. Group D compared the quantity-related bonus with the flat bonus. Group E compared the revenue-related bonus with the flat bonus. Group F: quantity-related compared to revenue-related bonus

Table 12. Summary of secondary outcomes findings

Studies	Patient or provider satisfaction	Equity	Comments on contextual factors influencing findings	Comments on interpretation of findings	Comments on data
Peabody 2010	Patient satisfaction improved for PBF group, but no change in insurance group	Not reported	It is possible that because the intervention was done in constrained economic setting, where the health of populations is consistently low, the marginal health benefits of better quality were higher	The similar benefits of the direct and the indirect incentives underscores the importance of considering system-level effects that may drive individual behaviour. While the study detected short-term health benefits after discharge, further examination would be needed to elucidate the longer-term benefits	Results assembled from a variety of papers, some published, some in draft. No subgroup analysis for those admitted with diarrhoea versus those admitted with pneumonia, which might be expected to influence the results. Authors only report on 2 indicators with significant improvement in abstract.
Basinga 2010	Not reported	Not reported	Preventive visit for child included immunisation, vitamin A, growth monitoring and distribution of mosquito nets. Some services high overall uptake (95% of women get some ANC), while others lower (less than 35% of women have institutional deliveries). Vaccination rates already at 65% to start with, so harder to raise. For quality of care, 70% of women got TT injections at ANC, but only 46% of overall activities which were prescribed were carried out by providers at ANC visits.	The authors conclude that P4P had the greatest impact on services which provided most reward and were most under the control of providers. Some, like tetanus vaccination during ANC, are not only funded but also allow providers to lobby women to deliver there and also increase their quality score. ANC too poorly funded.	Analysis controls for household characteristics (including enrolment in health insurance), as well as assets. Actual figures are not reported. The authors report baseline levels (proportions) and effect estimates (probability increase) from multivariate regression models. Original data was not provided when requested
Soeters 2009	Quality of care as perceived by households declined for both groups, with no significant difference between the two	Out of pocket health expenditures reported (as proportion of overall per capita income, and for the	Huge baseline difference for institutional deliveries, and also for other measures. Two different donor support models in operation in intervention and control areas. In 2006 free care was announced nationally for under-5s and deliveries - this will also have impacted on the variables studied. Provinces received different amounts of funding from donors and for free care ($4.3 per head per year in the intervention provinces; $4.7 in one control; $7.7 in another, at least for 2007). The annual per capita government assistance to health centres in the PBF and control provinces approximately doubled between 2006 and 2008 but it remained at the still low level of $0.19 to $0.36 per capita per year. Assistance also provided in kind, which not quantified for either group.		Original data provided by author

Table 12. Summary of secondary outcomes findings *(Continued)*

		poorest 25%), but only aggregate data presented			
Canavan 2008	No significant difference in user satisfaction ratings between PBF and non-PBF facilities, though mission facilities score slightly higher on quality of care ratings (but absence of baseline data). The fact that the contracts were signed between Cordaid and the diocese, albeit maintaining the purchaser-provider split essential for PBF, proved to be a major disadvantage in instilling responsibility for results and ownership of the performance indicators at health facility level which was often not involved or aware of contract negotiations and agreements but responsible for its results. The health facilities managers highlighted dissatisfaction with several of the indicators selected by Cordaid and the corresponding targets set. Overall staff satisfaction was found to be similar across PBF and non-PBF facilities. Intrinsic motivation factors emerged as most important	Not reported	Other relevant context information include that coverage calculations not be accurate as patients have been moving over to government hospitals where treatment is affordable; there is also inappropriate referral due to poorly functioning primary care facilities; there has been an increase in VCT services due to nationwide campaigns; users from outside the catchment areas are attending government health facilities; there has been construction of new dispensaries and subsequent division of catchment population by district health councils, as well as movement of qualified staff, lack of qualified staff and thus users travelling to other nearby health facilities.	Lack of involvement of different stakeholders in the conceptualisation and institutional set up of PBF has proven a weakness in the programme, and could affect its sustainability. Authors comment that additional financial and technical support are needed to make PBF operational. Other preconditions for success are also missing (e.g. inadequate staffing numbers). Tanzania has a 62% shortfall in staffing, relative to MoH norms.	See context - doubts on catchment denominators make all coverage figures questionable. Original data obtained from authors
Vergeer 2008	Higher quality of care scores by patients for PBF facilities (but no baseline with which to compare it). Patients generally satisfied but some concerns for both groups. Health workers face	Not established, but authors note that no emphasis with-	Mission facilities are funded by government, but there is not complete parity (e.g. pay is lower, exacerbating staff shortages). The facility revenues were significantly higher for mission facilities, compared to government ones, which will af-	Low awareness by staff of how PBF is meant to function - treated as another form of input financing. No link between payments and staff behaviour, except in one hospital. Authors note that the preconditions for effectiveness of PBF were not in	The denominator (per 1000 people) of the hospital delivery and VCT user rates both experienced problems due to concerns about the accuracy of the

Table 12. Summary of secondary outcomes findings *(Continued)*

	increased workload. Some complained about higher bonus payments for doctors - demotivating for those receiving less.	in PBF scheme on propoor measures (e.g. no focus on remote areas, which are more in need)	fect their ability to reach the targets. Health facilities also emphasised the significant attention already paid to increasing VCT consultations through other donor supported programmes. Attribution of the results solely to the Cordaid implemented PBF approach was difficult due to several confounding factors, with the most important one being the abolishment of user fees in April 2006, which also led to an increase in the utilisation of health facilities.	place, including adequate staffing, reliable information systems, quality assurance mechanisms, and systems for community involvement and feedback. Call for more context-specific indicators and funding	catchment population assigned to each facility, and led to facilities reporting different rates to the MoH and Cordaid. Original data requested from authors but are missing.
Soeters 2005	Many of the report's findings come from a survey of 64 health workers. However, there is no baseline so hard to attribute findings to intervention. Comments reflect on relationship with the managing NGOs. Staff in PBF areas are satisfied though question amounts paid for targets.	Not reported	Some differences which may affect results, e.g. health worker pay on average 30% higher in intervention provinces (at end, at least - no baseline data). Overall funding and support to each of the four provinces also different		The findings for the main outcome indicators reported come from the routine HMIS - risk of bias. Original data obtained from author.
Soeters 2008	Patient-assessed quality higher in PBF districts for most indicators (before and after - 81% satisfaction before for intervention, compared to 79% for controls, and 83% after, compared to 67% for controls). Significant improvement in patient assessments of quality, drug availability and composite quality score (as assessed by patients) in PBF areas, compared to controls. Managers said to be dissatisfied in most facilities, particularly with the level of external support	Household payments as a proportion of income by the poorest reduced more in intervention districts (by 63.5%, compared to 21.9% reduction in controls), though at much high-	The DRC is a particular 'failed state' context, where facilities support themselves largely without government support. Each district had different previous investments and donor support. There is no evidence presented that they were similar at baseline - in fact, there are clear differences presented (e.g. higher household contributions before in the PBF districts). Although difference-in-difference in theory allows for these starting differences, it does not fully allow for ceiling effects etc. Authors comment that the $0.30 per capita for hospital care was too low in retrospect, and that FP coverage is still very low		Some of the authors' conclusions are not supported by the data. E.g. "we extrapolate that a contribution of approximately 7-10% of household income for health would be acceptable in the DRC". But there is no investigation of acceptability in the paper. There is also a mismatch between findings in the paper and the very positive conclusions in the abstract.

Table 12. Summary of secondary outcomes findings *(Continued)*

		er level to start with. For the poor (second quartile), it is the reverse, with a 76.5% reduction in the controls, compared to a 36.2% reduction in intervention district households		
Liu 2003	Not reported	Not reported	The study took place in the context of high hospital autonomy and increasing reliance on revenues from patients (only 6% of hospital revenue coming from government at the time). This led to strong pressures to increase facility revenues, and hence the growth of bonus schemes to motivate staff, changing from flat to quantity-related to revenue-related, over time. Despite some prices being fixed by the government, hospital revenues grew by nearly 19% per year over the period (over 11% per year in real terms)	Prices were adjusted to allow for inflation (1975 as base year). Data obtained from authors. Interventions were introduced at different time points over 22 year period.
Quy 2003	Not reported	Not reported	In Vietnam there was a low detection rate - around 60% - hence focus on including private practitioners. However, 58% of those detected in this trial defaulted from treatment, indicating the importance of the next stage as well. An estimated 40% of TB patients seek care in the private sector in Vietnam. Cure rates are poor - less than 50%. Free treatment was available from the NTP.	Results re-analysed using ITS. Adjustment for seasonality significantly affects results. Short-time frame noted.

Note on abbreviations: ANC = antenatal care; DRC = Democratic Republic of Congo; HMIS = Health Management Information System; ITS = interrupted time series; MoH = Ministry of Health; NGO = non-governmental organisation; NTP = National Tuberculosis Programme; P4P = payment for performance; PBF = performance-based financing; TB = tuberculosis; TT = tetanus toxoid; VCT = Voluntary counselling and testing

Table 13. Studies identified, by type, level of targeting, incentive size and presence of additional ancillary components

Target	Number of studies of paying for performance schemes including ancillary components	Number of studies of paying for performance without additional (only in intervention group) ancillary components

Table 13. Studies identified, by type, level of targeting, incentive size and presence of additional ancillary components *(Continued)*

	Substantial incentives (> 20% of total funding)	Small incentives (20% of total funding or less)	Substantial incentives (> 20% of total funding)	Small incentives (20% of total funding or less)
Individuals	Quy 2003 (magnitude unknown)		Liu 2003 (magnitude unknown)	
		Peabody 2010		
Facilities	Soeters 2009; Soeters 2005; Soeters 2008	Canavan 2008; Vergeer 2008	Basinga 2010	0
Areas/sub-national organisations	0	0	0	0
National governments	0	0	0	0
Multiple	0	0	0	0

Table 14. Study characteristics of included studies

Studies	Allocation to study	Proportion of eligible providers participating in evaluation	Power calculation	Design	Comments
Peabody 2010	30 hospitals, randomly allocated to 3 groups of 10. Participation in scheme voluntary. Groups matched based on supply and demand characteristics.	Not clear	Not done	Cluster-randomised controlled trial	Randomised allocation but assessment of outcomes not blinded. Physician and facility characteristics compared at start and end and no significant differences found between groups. At baseline, there were 501 patients in the control arm and 496 in the intervention arm; at round 2, there were 560 and 596 patients respectively. 89% response rate by parents
Basinga 2010	Allocation by district (8 + 8). 80 facilities in intervention group and 86 in control. Random sample of 2158 households (baseline and 24 months later). 620 women who had delivered surveyed in intervention areas and 870 in control.	Not clear - may be 100%	Done	Designed prospectively as CRCT but with some adjustments made to allocation for practical reasons	88% of households reinterviewed (12% drop-out similar for intervention and control). Baseline study confirmed no significant differences between intervention and control facilities, except for 4 or more ANC visits, which women in intervention areas were significantly more likely to make at baseline
Soeters 2009	29 facilities surveyed (total for both groups). Household survey of 500, repeated before (2006) and after (2008)	40 facilities participated in the PBF.	Not done	CBA, designed prospectively	Considerable differences noted between intervention and control groups. Selection of facilities not clearly described. No information on drop-out rates for household survey. Not clear how the choice of indicators to include in the

Table 14. Study characteristics of included studies *(Continued)*

		So it appears that around a third of these were surveyed for the study			household survey was made (not all performance areas are covered).
Canavan 2008	12 facilities included: 1 hospital, health centre and dispensary for each area - 2 dioceses (intervention areas) and government 'controls' included	We calculate about one-tenth of mission-supported facilities	Not done	CBA, designed retrospectively so no baseline data for intervention and control sites	Considerable differences noted between intervention and control groups. Selection of facilities not clearly described
Vergeer 2008	8 facilities included: 6 hospitals (4 in intervention and two in controls) and 2 health centres (1 intervention, 1 control)	We calculate about 8% of mission-supported facilities; not clear for government ones	Not done	CBA, designed retrospectively so no baseline data for intervention and control sites	Considerable differences noted between intervention and control groups. Some facilities chosen 'at random' and others for convenience. Small sample.
Soeters 2005	16 health centres across 4 provinces of Rwanda (2 intervention and 2 control provinces); in each province 2 were government-owned and 2 church-owned	15%	Not done	CBA, designed retrospectively so no baseline data for intervention and control sites	Different funding mechanisms to each province. No baseline data. Study relied on retrospective analysis of routine data collected by facilities (HMIS). Health centres said to be chosen at random but no data presented on comparability. Authors of study also involved in design of intervention
Soeters 2008	Study methods included pre/post household survey, professional review of 12 health centres in PBF districts, compared with 10 in controls, and semi-structured interviews with 22 health centre managers	Roughly one-third of participating health centres in intervention districts; proportion not clear for controls	Not done	CBA, designed prospectively, with survey of 240 households in PBF districts and 200 in controls in 2005; repeated in 2008	There is no description of how the facilities were selected. The authors note that the baseline household survey was carried out in November 2005, while the evaluation was conducted in February 2008. Seasonal differences may, therefore, have affected the results. Another potential bias was the 85% increase in the per capita annual cash income in the overall study area: from $65 in 2005 to $122 in 2008. Due to lack of financial and human resources, a relatively small sample of 440 households was used, contributing to some of the findings being statistically weak and further compromised by the cluster survey design. Finally, most indicators were selected before the intervention, while some were selected during the analysis. Some indicators, like TB and vitamin A, despite being within

Table 14. Study characteristics of included studies *(Continued)*

					the funded indicators, were not included in the survey - it is not clear why
Liu 2003	Natural experiment, benefiting from adoption of different bonus schemes by hospitals over time	85% of providers (108 out of 127 hospitals approached) gave 'valid answers'	Not done	ITS, using retrospective data from 1975 to 1997, as well as cross-sectional data for 1997	Changes introduced at different time points. Number of points not uniform across all groups. Different numbers of hospitals per group. No clear if there is response bias (15% non-response). Authors do not use formal ITS techniques; trend analysis appears to be visual.
Quy 2003	All formally registered private practitioners were invited to join in the intervention districts. No information is given on how the 2 districts were selected	52% of doctors and 100% of pharmacists	Not done	Presented as CBA, but with only 1 intervention and control group. Has been reanalysed as an ITS.	Complex intervention (not just incentive payments), which limits attribution of effect to payments alone. Follow-up also for very short period (1 year)

ANC: antenatal care; CBA = controlled before-after study; CRCT = cluster-randomised controlled trial; ITS = interrupted time series study; PBF = performance-based financing;TB = tuberculosis

APPENDICES

Appendix 1. Search strategies

CENTRAL

#1 MeSH descriptor Employee Incentive Plans, this term only

#2 MeSH descriptor Physician Incentive Plans, this term only

#3 MeSH descriptor Reimbursement, Incentive, this term only

#4 MeSH descriptor Fee-for-Service Plans, this term only

#5 MeSH descriptor Contract Services, this term only

#6 pay* NEAR/3 performance:ti,ab

#7 "p4p":ti,ab

#8 pay* NEAR/3 quality:ti,ab

#9 (fee NEAR/3 service or fees NEAR/3 services or fee NEAR/3 services or fees NEAR/3 service):ti,ab

#10 (incentive* or compensatory or reimbursement) NEAR plan*:ti,ab

#11 contract NEXT service*:ti,ab

#12 result? NEXT based:ti,ab

#13 performance NEXT based:ti,ab

#14 (result? or performance or output or "out put") NEAR/2 (financ* or fund* or pay* or disburs* or fee? or incentive? or initiative? or contract? or aid)

#15 (pay* or monetary or economic or financial or reimbursement) NEXT incentive?:ti,ab

#16 (economic or financial) NEXT (reward? or bonus$):ti,ab

#17 target* NEXT pay*:ti,ab

#18 MeSH descriptor Developing Countries, this term only

#19 MeSH descriptor Medically Underserved Area, this term only

#20 MeSH descriptor Africa explode all trees

#21 MeSH descriptor Asia explode all trees

#22 MeSH descriptor South America explode all trees

#23 MeSH descriptor Central America explode all trees

#24 MeSH descriptor Latin America, this term only

#25 ("American Samoa" or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or "Costa Rica" or Croatia or Dominica or "Equatorial Guinea" or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or "South Africa" or "Saint Kitts and Nevis" or "Saint Lucia" or "Saint Vincent and the Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia or Guinea or Libia or libyan or Mayotte or "Northern Mariana Islands" or "Russian Federation" or Samoa or Serbia or "Slovak Republic" or "St Kitts and Nevis" or "St Lucia" or "St Vincent and the Grenadines"):ti,ab,kw

#26 (Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or "Dominican Republic" or Ecuador or Egypt or "El Salvador" or Fiji or "Georgia (Republic)" or Guam or Guatemala or Guyana or Honduras or "Indian Ocean Islands" or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or "Marshall Islands" or Micronesia or "Middle East" or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or "Sri Lanka" or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu or Bosnia or "Cape Verde" or Gaza or Georgia or Kiribati or Macedonia or Maldives or "Marshall Islands" or Palestine or "Syrian Arab Republic" or "West Bank"):ti,ab,kw

#27 (Afghanistan or Bangladesh or Benin or "Burkina Faso" or Burundi or Cambodia or "Central African Republic" or Chad or Comoros or "Democratic Republic of the Congo" or "Cote d'Ivoire" or Eritrea or Ethiopia or Gambia or Ghana or Guinea or "Guinea-Bissau" or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or "Papua New Guinea" or Rwanda or Senegal or "Sierra Leone" or Somalia or Sudan or Tajikistan or Tanzania or "East Timor" or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe or Burma or Congo or Kyrgyz or Lao or "North Korea" or "Salomon Islands" or "Sao Tome" or Timor or "Viet Nam"):ti,ab,kw

#28 (developing or less* NEXT developed or "third world" or "under developed" or "middle income" or low* NEXT income or underserved or "under served" or deprived or poor*) NEXT (count* or nation? or state? or population?):ti,ab,kw

#29 (Africa or Asia or "South America" or "Latin America" or "Central America" or lmic or lmics):ti,ab,kw

#30 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17)

#31 (#18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29)

#32 (#30 AND #31)

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE

1 Employee Incentive Plans/ (1441)

2 Physician Incentive Plans/ (1860)

3 Reimbursement, Incentive/ (2186)

4 Fee-for-Service Plans/ (2192)

5 Contract Services/ (9605)

6 (conditional adj3 (pay$ or transfer?)).tw. (81)

7 (pay$ adj3 performance).tw. (1189)

8 p4p.tw. (246)

9 (pay$ adj3 quality).tw. (410)

10 fee? for service?.tw. (2983)

11 ((incentive? or compensatory or reimbursement) adj plan?).tw. (189)

12 contract service?.tw. (89)

13 result? based.tw. (5324)

14 performance based.tw. (1864)

15 ((result? or performance or output or out put) adj2 (financ$ or fund$ or pay$ or disburs$ or fee? or incentive? or initiative? or contract? or aid)).tw. (5198)

16 payment incentive?.tw. (76)

17 monetary incentive?.tw. (398)

18 economic incentive?.tw. (577)

19 financial incentive?.tw. (1896)

20 reimbursement incentive?.tw. (41)

21 ((economic or financial) adj (reward? or bonus$)).tw. (320)

22 target$ pay$.tw. (28)

23 or/1-22 (32884)

24 Developing Countries/ (55903)

25 Medically Underserved Area/ (4841)

26 exp Africa/ or exp "Africa South of the Sahara"/ or exp Asia/ or exp South America/ or exp Latin America/ or exp Central America/ (629188)

27 (Africa or Asia or South America or Latin America or Central America).tw. (78504)

28 (American Samoa or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or Costa Rica or Croatia or Dominica or Equatorial Guinea or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or South Africa or "Saint Kitts and Nevis" or Saint Lucia or "Saint Vincent and the Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia).sh,tw. or Guinea.tw. or Libia.tw. or libyan.tw. or Mayotte.tw. or Northern Mariana Islands.tw. or Russian Federation.tw. or Samoa.tw. or Serbia.tw. or Slovak Republic.tw. or "St Kitts and Nevis".tw. or St Lucia.tw. or "St Vincent and the Grenadines".tw. (402675)

29 (Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or Dominican Republic or Ecuador or Egypt or El Salvador or Fiji or "Georgia (Republic)" or Guam or Guatemala or Guyana or Honduras or Indian Ocean Islands or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or Marshall Islands or Micronesia or Middle East or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or Sri Lanka or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu).sh,tw. or Bosnia.tw. or Cape Verde.tw. or Gaza.tw. or Georgia.tw. or Kiribati.tw. or Macedonia.tw. or Maldives.tw. or Marshall Islands.tw. or Palestine.tw. or Syrian Arab Republic.tw. or West Bank.tw. (231749)

30 (Afghanistan or Bangladesh or Benin or Burkina Faso or Burundi or Cambodia or Central African Republic or Chad or Comoros or "Democratic Republic of the Congo" or Cote d'Ivoire or Eritrea or Ethiopia or Gambia or Ghana or Guinea or Guinea-Bissau or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or Papua New Guinea or Rwanda or Senegal or Sierra Leone or Somalia or Sudan or Tajikistan or Tanzania or East Timor or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe).sh,tw. or Burma.tw. or Congo.tw. or Kyrgyz.tw. or Lao.tw. or North Korea.tw. or Salomon Islands.tw. or Sao Tome.tw. or Timor.tw. or Viet Nam.tw. (335390)

31 ((developing or less$ developed or third world or under developed or middle income or low income or underserved or under served or deprived or poor$) adj (count$ or nation? or state? or population?)).tw. (40286)

32 (lmic or lmics).tw. (114)

33 or/24-32 (1061726)

34 randomized controlled trial.pt. (309565)

35 random$.tw. (553751)

36 intervention$.tw. (407857)

37 control$.tw. (2124086)

38 evaluat$.tw. (1734641)

39 effect$.tw. (3909417)

40 or/34-39 (6474912)

41 Animals/ (4769239)

42 Humans/ (11785790)

43 41 not (41 and 42) (3520949)

44 40 not 43 (4832589)

45 23 and 33 and 44 (1001)

46 45 (1001)

47 limit 46 to yr="2009 -Current" (283)

EMBASE (Ovid)

1. (pay$ adj3 performance).tw.

2. p4p.tw.

3. (pay$ adj3 quality).tw.

4. fee? for service?.tw.

5. ((incentive? or compensatory or reimbursement) adj plan?).tw.

6. contract service?.tw.

7. result? based.tw.

8. performance based.tw.

9. ((result? or performance or output or out put) adj2 (financ$ or fund$ or pay$ or disburs$ or fee? or incentive? or initiative? or contract? or aid)).tw.

10. payment incentive?.tw.

11. monetary incentive?.tw.

12. economic incentive?.tw.

13. financial incentive?.tw.

14. reimbursement incentive?.tw.

15. ((economic or financial) adj (reward? or bonus$)).tw.

16. target$ pay$.tw.

17. or/1-16

18. Developing Country/

19. exp Africa/ or exp Asia/ or exp "South and Central America"/

20. (Africa or Asia or South America or Latin America or Central America).tw.

21. (American Samoa or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or Costa Rica or Croatia or Dominica or Equatorial Guinea or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or South Africa or "Saint Kitts and Nevis" or Saint Lucia or "Saint Vincent and the Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia or Guinea or Libia or libyan or Mayotte or Northern Mariana Islands or Russian Federation or Samoa or Serbia or Slovak Republic or "St Kitts and Nevis" or St Lucia or "St Vincent and the Grenadines").sh,tw.

22. (Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or Dominican Republic or Ecuador or Egypt or El Salvador or Fiji or "Georgia (Republic)" or Guam or Guatemala or Guyana or Honduras or Indian Ocean Islands or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or Marshall Islands or Micronesia or Middle East or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or Sri Lanka or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu or Bosnia or Cape Verde or Gaza or Georgia or Kiribati or Macedonia or Maldives or Marshall Islands or Palestine or Syrian Arab Republic or West Bank).sh,tw.

23. (Afghanistan or Bangladesh or Benin or Burkina Faso or Burundi or Cambodia or Central African Republic or Chad or Comoros or "Democratic Republic of the Congo" or Cote d'Ivoire or Eritrea or Ethiopia or Gambia or Ghana or Guinea or Guinea-Bissau or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or Papua New Guinea or Rwanda or Senegal or Sierra Leone or Somalia or Sudan or Tajikistan or Tanzania or East Timor or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe or Burma or Congo or Kyrgyz or Lao or North Korea or Salomon Islands or Sao Tome or Timor or Viet Nam).sh,tw.

24. ((developing or less$ developed or third world or under developed or middle income or low income or underserved or under served or deprived or poor$) adj (count$ or nation? or state? or population?)).tw.

25. (lmic or lmics).tw.

26. or/18-25

27. Randomized Controlled Trial/

28. Time Series Analysis/

29. random$.tw.

30. experiment$.tw.

31. control$.tw.

32. (time adj series).tw.

33. (pre test or pretest or post test or posttest).tw.

34. impact.tw.

35. intervention$.tw.

36. chang$.tw.

37. evaluat$.tw.

38. effect?.tw.

39. compar$.tw.

40. or/27-39

41. nonhuman/

42. 40 not 41

43. 17 and 26 and 42

PsycINFO (Ovid)

1. Monetary Incentives/

2. Monetary Rewards/

3. Fee for Service/

4. Professional Fees/

5. "pay for performance".id.

6. (pay$ adj3 performance).tw.

7. p4p.tw.

8. (pay$ adj3 quality).tw.

9. fee? for service?.tw.

10. ((incentive? or compensatory or reimbursement) adj plan?).tw.

11. contract service?.tw.

12. result? based.tw.

13. performance based.tw.

14. ((result? or performance or output or out put) adj2 (financ$ or fund$ or pay$ or disburs$ or fee? or incentive? or initiative? or contract? or aid)).tw.

15. payment incentive?.tw.

16. monetary incentive?.tw.

17. economic incentive?.tw.

18. financial incentive?.tw.

19. reimbursement incentive?.tw.

20. ((economic or financial) adj (reward? or bonus$)).tw.

21. target$ pay$.tw.

22. or/1-21

23. Developing Countries/

24. (Africa or Asia or South America or Latin America or Central America).id,tw.

25. (American Samoa or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or Costa Rica or Croatia or Dominica or Equatorial Guinea or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or South Africa or "Saint Kitts and Nevis" or Saint Lucia or "Saint Vincent and the Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia or Guinea or Libia or libyan or Mayotte or Northern Mariana Islands or Russian Federation or Samoa or Serbia or Slovak Republic or "St Kitts and Nevis" or St Lucia or "St Vincent and the Grenadines").tw.

26. (Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or Dominican Republic or Ecuador or Egypt or El Salvador or Fiji or "Georgia (Republic)"

or Guam or Guatemala or Guyana or Honduras or Indian Ocean Islands or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or Marshall Islands or Micronesia or Middle East or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or Sri Lanka or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu or Bosnia or Cape Verde or Gaza or Georgia or Kiribati or Macedonia or Maldives or Marshall Islands or Palestine or Syrian Arab Republic or West Bank).tw.

27. (Afghanistan or Bangladesh or Benin or Burkina Faso or Burundi or Cambodia or Central African Republic or Chad or Comoros or "Democratic Republic of the Congo" or Cote d'Ivoire or Eritrea or Ethiopia or Gambia or Ghana or Guinea or Guinea-Bissau or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or Papua New Guinea or Rwanda or Senegal or Sierra Leone or Somalia or Sudan or Tajikistan or Tanzania or East Timor or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe or Burma or Congo or Kyrgyz or Lao or North Korea or Salomon Islands or Sao Tome or Timor or Viet Nam).tw.

28. ((developing or less$ developed or third world or under developed or middle income or low income or underserved or under served or deprived or poor$) adj (count$ or nation? or state? or population?)).tw.

29. (lmic or lmics).tw.

30. or/23-29

31. 22 and 30

32. Treatment Effectiveness Evaluation/

33. Clinical Trials/

34. random$.tw.

35. experiment$.tw.

36. control$.tw.

37. (time adj series).tw.

38. (pre test or pretest or post test or posttest).tw.

39. impact.tw.

40. intervention$.tw.

41. chang$.tw.

42. evaluat$.tw.

43. effect?.tw.

44. compar$.tw.

45. or/32-44

46. 31 and 45

47. limit 31 to ("0400 empirical study" or "0451 prospective study" or "2000 treatment outcome/randomized clinical trial")

48. 46 or 47

EconLit (Ovid)

1. Compensation Packages; Payment Methods.sh.

2. Personnel Management; executive compensation.sh.

3. "Personnel Economics: Compensation and Compensation Methods and Their Effects".sh.

4. Contingent Pricing; Futures Pricing; option pricing.sh.

5. Payout Policy.sh.

6. payment systems.kw.

7. pay-for-performance.kw.

8. compensation.kw.

9. incentives.kw.

10. performance related pay.kw.

11. executive compensation.kw.

12. bonuses.kw.

13. (pay$ adj3 performance).tw.

14. p4p.tw.

15. (pay$ adj3 quality).tw.

16. fee? for service?.tw.

17. ((incentive? or compensatory or reimbursement) adj plan?).tw.

18. contract service?.tw.

19. result? based.tw.

20. performance based.tw.

21. ((result? or performance or output or out put) adj2 (financ$ or fund$ or pay$ or disburs$ or fee? or incentive? or initiative? or contract? or aid)).tw.

22. ((pay$ or monetary or economic or financial or reimbursement) adj incentive?).tw.

23. incentive pay.tw.

24. ((monetary or economic or financial) adj (reward? or bonus$)).tw.

25. target$ pay$.tw.

26. or/1-25

27. Health: General.sh.

28. Health: Other.sh.

29. Health: Government Policy; Regulation; Public Health.sh.

30. "National Government Expenditures and Health".sh.

31. Analysis of Health Care Markets.sh.

32. Health Production.sh.

33. health.kw.

34. health care.kw.

35. healthcare.kw.

36. medical care.kw.

37. hospital.kw.

38. hospitals.kw.

39. health insurance.kw.

40. managed care.kw.

41. doctor.kw.

42. doctors.kw.

43. physician.kw.

44. physicians.kw.

45. (health$ or medical$).tw.

46. or/27-45

47. 26 and 46

48. Developing Countries.hw.

49. (Africa or Asia or South America or Latin America or Central America).kw,tw.

50. (American Samoa or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or Costa Rica or Croatia or Dominica or Equatorial Guinea or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or South Africa or "Saint Kitts and Nevis" or Saint Lucia or "Saint Vincent and the Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia or Guinea or Libia or libyan or Mayotte or Northern Mariana Islands or Russian Federation or Samoa or Serbia or Slovak Republic or "St Kitts and Nevis" or St Lucia or "St Vincent and the Grenadines").kw,tw.

51. (Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or Dominican Republic or Ecuador or Egypt or El Salvador or Fiji or "Georgia (Republic)" or Guam or Guatemala or Guyana or Honduras or Indian Ocean Islands or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or Marshall Islands or Micronesia or Middle East or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or Sri Lanka or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu or Bosnia or Cape Verde or Gaza or Georgia or Kiribati or Macedonia or Maldives or Marshall Islands or Palestine or Syrian Arab Republic or West Bank).kw,tw.

52. (Afghanistan or Bangladesh or Benin or Burkina Faso or Burundi or Cambodia or Central African Republic or Chad or Comoros or "Democratic Republic of the Congo" or Cote d'Ivoire or Eritrea or Ethiopia or Gambia or Ghana or Guinea or Guinea-Bissau or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or Papua New Guinea or Rwanda or Senegal or Sierra Leone or Somalia or Sudan or Tajikistan or Tanzania or East Timor or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe or Burma or Congo or Kyrgyz or Lao or North Korea or Salomon Islands or Sao Tome or Timor or Viet Nam).kw,tw.

53. ((developing or less$ developed or third world or under developed or middle income or low income or underserved or under served or deprived or poor$) adj (count$ or nation? or state? or population?)).kw,tw.

54. (lmic or lmics).kw,tw.

55. or/48-54

56. 47 and 55

Sociological Abstracts and Social Services Abstracts (CSA)

(((kw=pay* within 3 performance) or(kw=p4p) or(kw=pay* within 3 quality) or(kw=fee within 3 service*) or(kw=fees within 3 service*) or(kw=(incentive* or compensatory or reimbursement) within 1 plan*) or(kw=contract service*) or(kw=result* based) or(kw=performance based) or(kw=pay* incentive*) or(kw=monetary incentive*) or(kw=economic incentive*) or(kw=financial incentive*) or(kw=reimbursement incentive*) or(kw=economic reward*) or(kw=economic bonus*) or(kw=financial reward*) or(kw=financial bonus*) or(kw=monetary reward*) or(kw=monetary bonus*) or(kw=target* pay*)) and(KW=(health* or medic* or hospital* or primary care or physician* or doctor* or general practitioner* or nurse*))) and((KW=(developing or less* developed or third world or under developed or middle income or low income or underserved or under served or deprived or poor*) within 1 (count* or nation* or state* or population*)) or(KW=(africa or asia or southe america or latin america or central America or lmic or lmics)) or(kw=(American Samoa or Argentina or Belize or Botswana or Brazil or Bulgaria or Chile or Comoros or Costa Rica or Croatia or Dominica or Equatorial Guinea or Gabon or Grenada or Hungary or Kazakhstan or Latvia or Lebanon or Libya or Lithuania or Malaysia or Mauritius or Mexico or Micronesia or Montenegro or Oman or Palau or Panama or Poland or Romania or Russia or Seychelles or Slovakia or South Africa or "Saint Kitts and Nevis" or Saint Lucia or "Saint Vincent and the

Grenadines" or Turkey or Uruguay or Venezuela or Yugoslavia or Guinea or Libia or libyan or Mayotte or Northern Mariana Islands or Russian Federation or Samoa or Serbia or Slovak Republic or "St Kitts and Nevis" or St Lucia or "St Vincent and the Grenadines")) or(kw=(Albania or Algeria or Angola or Armenia or Azerbaijan or Belarus or Bhutan or Bolivia or "Bosnia and Herzegovina" or Cameroon or China or Colombia or Congo or Cuba or Djibouti or Dominican Republic or Ecuador or Egypt or El Salvador or Fiji or "Georgia (Republic)" or Guam or Guatemala or Guyana or Honduras or Indian Ocean Islands or Indonesia or Iran or Iraq or Jamaica or Jordan or Lesotho or "Macedonia (Republic)" or Marshall Islands or Micronesia or Middle East or Moldova or Morocco or Namibia or Nicaragua or Paraguay or Peru or Philippines or Samoa or Sri Lanka or Suriname or Swaziland or Syria or Thailand or Tonga or Tunisia or Turkmenistan or Ukraine or Vanuatu or Bosnia or Cape Verde or Gaza or Georgia or Kiribati or Macedonia or Maldives or Marshall Islands or Palestine or Syrian Arab Republic or West Bank)) or(kw=(Afghanistan or Bangladesh or Benin or Burkina Faso or Burundi or Cambodia or Central African Republic or Chad or Comoros or "Democratic Republic of the Congo" or "Cote d'Ivoire" or Eritrea or Ethiopia or Gambia or Ghana or Guinea or Guinea-Bissau or Haiti or India or Kenya or Korea or Kyrgyzstan or Laos or Liberia or Madagascar or Malawi or Mali or Mauritania or Melanesia or Mongolia or Mozambique or Myanmar or Nepal or Niger or Nigeria or Pakistan or Papua New Guinea or Rwanda or Senegal or Sierra Leone or Somalia or Sudan or Tajikistan or Tanzania or East Timor or Togo or Uganda or Uzbekistan or Vietnam or Yemen or Zambia or Zimbabwe or Burma or Congo or Kyrgyz or Lao or North Korea or Salomon Islands or Sao Tome or Timor or Viet Nam)))

LILACS

((Employee Incentive Plans) or (Physician Incentive Plans) or (Reimbursement, Incentive) or (Fee-for-Service Plans) or (Contract Services)) and (random$ or intervention$ or control$ or evaluat$ or effect$ or compare$ or chang$ or experiment$ or impact) (Subject descriptor)

((pay$ AND performance) or (p4p) or (monetary AND incentive$ AND performance) or (money AND incentive$ AND performance) or (economic AND incentive$ AND performance) or (financial AND incentive$ AND performance)) and (random$ or intervention$ or control$ or evaluat$ or effect$ or compare$ or chang$ or experiment$ or impact) (Words)

((monetary AND reward$ AND performance) or (money AND reward$ AND performance) or (economic AND reward$ AND performance) or (financial AND reward$ and performance) or (monetary AND bonus$ AND performance) or (money AND bonus$ AND performance) or (economic AND bonus$ AND performance) or (financial AND bonus$ AND performance) or (target$ AND pay$ AND performance) or (performance AND contract$)) and (random$ or intervention$ or control$ or evaluat$ or effect$ or compare$ or chang$ or experiment$ or impact) (Words)

WHOLIS

(pay$ NEAR3 perform$) or (p4p) or (pay$ NEAR3 quality) or (result? ADJ based) or (performance ADJ based) or (incentive ADJ plan?) or (compensat$ ADJ plan?) or (reimburs$ ADJ plan?) or (fee for service)

(pay$ NEAR3 incentive?) or (reimburs$ ADJ incentive?) or (monetary NEAR3 incentive?) or (money NEAR3 incentive?) or (economic NEAR3 incentive?) or (financial NEAR3 incentive?)

(economic NEAR3 reward?) or (financial NEAR3 reward?) or (monetary NEAR3 reward) or (economic NEAR3 bonus$) or (financial NEAR3 bonus$) or (monetary NEAR3 bonus$) or (target$ NEAR3 pay$) or (contract ADJ service?)

(result? NEAR3 financ$) or (result? NEAR3 fund$) or (result? NEAR3 pay$) or (result? NEAR3 disburs$) or (result? NEAR3 fee?) or (result? NEAR3 incentive?) or (result? NEAR3 initiative?) or (result? NEAR3 contract?) or (result? NEAR3 aid)

(performance NEAR3 financ$) or (performance NEAR3 fund$) or (performance NEAR3 pay$) or (performance NEAR3 disburs$) or (performance NEAR3 fee?)

(performance NEAR3 incentive?) or (performance NEAR3 initiative?) or (performance NEAR3 contract?) or (performance NEAR3 aid)

WHAT'S NEW

Date	Event	Description
1 May 2013	Amended	Minor edit, fixed the extra space at the end of the title

HISTORY

Protocol first published: Issue 3, 2009

Review first published: Issue 2, 2012

Date	Event	Description
13 February 2012	Amended	Minor edits

CONTRIBUTIONS OF AUTHORS

SW and AF drafted the protocol. SW, AF and FK developed the search strategy. FK and AKL reviewed and commented on drafts of the protocol. SW, FK, AF and AKL selected the studies and undertook data extraction. SW led in the drafting of the review, with the support of AF. All authors have reviewed and commented on the final draft.

DECLARATIONS OF INTEREST

AF was a co-author in a previous study on the topic of results-based financing (Oxman 2008).

SOURCES OF SUPPORT

Internal sources

- Norwegian Knowledge Centre for the Health Services, Norway.
- Immpact, University of Aberdeen, UK.

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol we stated that the primary analysis would be for comparisons where there were no important differences regarding ancillary components. However, as this applied only to one study (Basinga 2010), we did not follow this in the final review, presenting all findings from eligible studies (where data were available), but noting the constraints.

INDEX TERMS

Medical Subject Headings (MeSH)

*Developing Countries; *Reimbursement, Incentive; Quality Improvement [*economics] [standards]; Quality of Health Care [economics] [standards]

MeSH check words

Humans